1-1-2005

Functional analysis of maize GBSSI and SSI : characterization of recombinant chimeric proteins

Rachel Carol Huegel
Iowa State University

Functional analysis of maize GBSSI and SSI: characterization of recombinant chimeric proteins

by

Rachel Carol Huegel

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

Major: Plant Physiology

Program of Study Committee:
Alan Myers, Co-major Professor
Martha James, Co-major Professor
Mark Westgate
Peter Keeling

Iowa State University
Ames, Iowa
2005

Graduate College
Iowa State University

This is to certify that the master's thesis of

Rachel Carol Huegel

has met the requirements of Iowa State University

Signatures have been redacted for privacy

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

APPENDIX B

APPENDIX C

ACKNOWLEDGEMENTS

There are many people that have contributed to my education and research in which I feel the need to acknowledge. First, I would very much like to thank my major professors, Dr. Martha G. James and Dr. Alan M. Myers, for all of their guidance and expertise that they provided to bring me through this my education and project with success. They took me under their wings when I needed it most for which I am very grateful.

I would also like to thank Dr. Mark Westgate for his advice during the development of my career at Iowa State. I would also like to show appreciation to Dr. Westgate and Dr. Peter Keeling for their contributions and support as a part of my committee. My gratitude goes out to Dr. Keeling and Dr. Padma Commuri as they were the first people who introduced me to realm starch biosynthesis in which this entire project is based. I would also like to thank my lab mates for guiding me down the right path every time I hit a fork in the road.

Finally, I would like to thank my family and friends for their love, support, and understanding throughout my education and my life. Their influence has provided me with the determination and confidence needed to travel this incredible journey. Most importantly, I would like thank my husband for his steadfast love and unwavering encouragement especially during the times I felt like giving up.

ABSTRACT

Starch is a semi-crystalline glucose polymer synthesized in higher plants for the efficient storage of a large amount of glucose in plastids without osmotic consequences. Amylose and amylopectin are two different glucose polymers that make up the starch granule which are created from adenosine 5' diphosphate glucose (ADPG) via starch synthases (SS's), branching enzymes (BE's), debranching enzymes (DBE's), starch phosphorylase (SP), and disproportionating enzyme (D-enzyme).

This focus of this research is to learn more about SS's exact structure and function. There are five known SS's in maize including granule bound starch synthase I (GBSSI), starch synthase I (SSI), starch synthase IIa (SSIIa), starch synthase IIb (SSIIb), and starch synthase III (SSIII). GBSS is primarily involved in the synthesis of amylose and the remaining SS's are involved in the synthesis of amylopectin. The structure of each SS is similar to one another and consists of the glucan association domain (GLASS) and the glycosyl transferase domain (GLYTR) with a linker region connecting the two domains. The role of the GLASS domain is suggested to possess binding properties to glucan chains and the GLYTR domain is thought to be involved in the interaction with the glucan during the process of starch synthesis and glucan chain elongation.

This report investigates the how the interaction between the GLASS and GLYTR domains effects the degree of functionality of the enzyme by combining the GLASS and linker domains of GBSS with the GLYTR domain of SSI. The results indicate that the relationship between these two domains is involved in how the enzyme functions. The susceptibility of the linker region is also investigated to determine the consequences of recombination in this region. The results indicate that this domain can not withstand minor alteration without changing its enzymatic properties indicating its importance in SS enzymes. However, further research is required to explain the precise function of each SS domain in order to understand how exactly how each individual SS enzyme functions.

CHAPTER 1. INTRODUCTION

Rationale and significance

Starch biosynthesis has been investigated for a number of years and much is known about the enzymes that participate in starch synthesis. However, the exact mechanism in which the semi-crystalline starch granule is created *in vivo* remains a mystery. In order to understand how starch is synthesized, the exact structure and function of the enzymes involved in generating this glucan structure must be further investigated. This report examines starch synthase (SS) enzyme structure and function to provide additional insight into starch synthesis.

There are five known SSs in maize including granule bound starch synthase I (GBSSI), starch synthase I (SSI), starch synthase IIa (SSIIa), starch synthase IIb (SSIIb), and starch synthase III (SSIII). GBSS is primarily involved in the synthesis of amylose and the remaining SSs are involved in the synthesis of amylopectin. The structure of each SS is similar to one another and consists of the glucan association domain (GLASS) and the glycosyl transferase domain (GLYTR) with a linker region connecting the two domains (see Figure 1). The GLASS domain is thought possesses binding properties to glucan chains and the GLYTR domain suggested to be involved in the interaction with the glucan during the process of starch synthesis and glucan chain elongation (see Figure 2).

Alteration of SSs function in plants could result in a starch product possessing desirable properties. There are a number of industrial applications for starch besides human diets and livestock feed; for example, starch is a component in pharmaceuticals, film coating, packing material, cosmetics, and other various materials. If SSs can be altered then starch with lower viscosity temperatures or less retro-gradation may be created which would be of interest to a variety of manufacturers.

Another important direction of this research would be to use the binding domain of granule bound starch synthase (GBSS), one of the five SSs found in maize, and fuse it with a protein of value. This fusion protein would be entrapped within the starch granule providing

protection to the protein of value. Starch storing crops with added nutrition due to fusion proteins in the starch granule would be very applicable in livestock feed and perhaps even in the human diet.

This report investigates the how the interaction between the GLASS and GLYTR domains effects the degree of functionality of the enzyme by combining the GLASS and linker domains of GBSS with the GLYTR domain of SSI (see Figure 3). The results indicate that the relationship between these two domains is involved in how the enzyme functions. The susceptibility of the linker region is also investigated to determine the consequences of recombination in this region. The results indicate that this domain can not withstand minor alteration without changing its enzymatic properties indicating its importance in SS enzymes. However, further research is required to explain the precise function of each SS domain in order to understand how each individual SS enzyme functions.

Literature Review

Starch is a semi-crystalline glucose polymer synthesized in higher plants for the efficient storage of a large amount of glucose in plastids without osmotic consequences. It is used as a raw material for many industrial applications in addition to use in human diets and livestock feed. Starch is a component in the manufacturing of textiles, plastic, paper, some building materials, glue, and is binder/disintegration agent in pharmaceutical tablets [1]. These examples demonstrate why it is so important to understand how starch is made in plants. Starch has been vigorously studied for many years however the exact mechanism of how starch is synthesized in plants is not fully understood. This review covers the current understanding of starch biosynthesis, with an emphasis on starch synthases.

Synthesis of starch occurs in the plastids where the type of starch created can vary based on the kind of plastid it is made in. Transient starch is produced in chloroplasts in leaves where it is degraded at night to provide carbon for non-photosynthetic metabolism. Starch made in the amyloplasts of seeds provides a long-term carbon store for the next

generation [2]. Amylose and amylopectin are the glucose polymers that comprise the starch granule. Amylose is predominantly linear with α-1,4 linked glucose units. Amylopectin is highly branched and is also composed of α-1,4 linked glucose units but contains approximately 5% α-1,6 linkages which are the branch points [3]. The semi-crystalline structure of the starch granule is possible due to the large amount of organization of amylopectin. Linear chain distribution is ordered and branch points are clustered together in granular amylopectin. Highly branched regions alternate with unbranched regions allowing the linear chains to align in parallel arrays of double helices [4]. This arrangement of amylopectin allows for dense packaging of a large amount of glucose.

The first committed step of starch biosynthesis in plants and bacteria begins with the formation of adenosine 5' diphosphate glucose (ADPG) via ADPG pyrophosphorylase (AGP). Bacterial AGP has a homotetrameric structure encoded by a single gene [5]. Plant AGP is a heterotetramer made up of two large subunits and two small subunits which are each encoded by different genes [6]. The conjecture of the function of these subunits is that the large subunit modifies the regulatory properties of the small subunit which is the catalytic unit [7]. AGP in amyloplasts is positively regulated by 3-phosphoglycerate and negatively regulated by inorganic phosphate [6] but its regulation may not be as important as in other green organs [4].

The location of AGP is not consistent between all higher plants. In cereal endosperm the enzyme is mostly extra-plastidial but plastidial in other tissues while in all other plants it is primarily plastidial [8]. AGP in maize has been found to be approximately 95% extra-plastidial and the ADPG is transported into the plastid [9] via BT1 transporter protein [10]. Some AGP activity is found within the plastid and can act on glucose-1-phosphate that is transported into the plastid to make ADPG [9].

Starch synthases

Once ADPG is in the plastid, SSs utilize it to elongate glucose chains via α-1, 4 linkages. There are five known SSs in maize including granule bound starch synthase

(GBSSI), starch synthase I (SSI), starch synthase IIa (SSIIa), starch synthase IIb (SSIIb), and starch synthase III (SSIII). Recently, starch synthase IV (SSIV) was discovered in rice and may be present in other plants [11]. GBSSI is primarily involved in the synthesis of amylose [12, 13], and the remaining SSs are involved in amylopectin synthesis [2]. Orthologous genes of SSs have very similar structure when comparing the exon-intron structures of all available genes encoding SSs from *A. tauschii*, barley, and *Arabidopsis* [14]. Dian *et al.* have found gene duplication of the SSs in rice, and suggest that there are two main groups; Group II genes synthesize early transient starch mainly in leaves, and Group I genes synthesize storage starch found mainly in the endosperm [11]. The structure of all five maize SSs is similar to glycogen and sucrose synthases, and is predicted to consist of two catalytic domains with a flexible linker region between the two to allow for open and closed formations of the active site [15]. The structural similarity between SSs, glycogen synthase, and sucrose synthase is not surprising as they are all closely related within families 4 and 5 of glycosyltransferases which facilitate the transfer of a monosaccharide from a donor to an acceptor [15].

A conserved amino acid sequence in plant *waxy* proteins of wheat, barley, maize, rice, potato, and pea of KTGGL located towards the N-terminus of the mature protein was also found in the glycogen synthase proteins of *E. coli* and animal muscle. The protein's function is implicated in the binding site for the substrates ADPG and UDP-glucose [16, 17]. Mason-Gamer *et al.* [17] discovered when the K in this region is changed in the GBSS enzyme a null allele occurs, indicating its importance in creating a functional enzyme. This same site is conserved between all SSs in higher plants and had been widely accepted as the putative substrate binding site of ADPG until Gao *et al.* found that this site on maize SSIIa was not directly involved in ADPG binding but instead alteration of this region affected primer preference [18]. Another structural similarity occurs between maize SSI, SSIIa, SSIIb, and SSIII, where there is an N-terminal extension known as a "flexible arm" located upstream from the KTGGL site; its exact function is yet to be determined [19].

Commuri *et al.* [20] showed that the structure of maize SSs were highly comparable to UDP-N-acetylglucosamine 2-epimerase, also a glycosyltransferase, from the results of protein threading on 3D-PSSM [21] (see Figure 1). They named the two domains of SSs glucan association domain (GLASS) and glycosyl transferase domain (GLYTR) because they envisaged the GLASS domain's purpose to be binding to a specific glucan, such as a growing amylopectin chain, while the GLYTR domain would function in the interaction of the addition of a glucose unit to the growing glucan chain (see Figure 2). Defining SS domain functions explains in more detail exactly how SSs elongate glucose chains.

Granule bound starch synthase

GBSSI is somewhat different from the other SSs in that its primary function is to elongate amylose in endosperm and is encoded by the *Waxy* (*Wx*) locus in maize (and other cereals such as rice, barley, and wheat) [12, 13]. It is the major starch granule bound protein in maize; its tissue-specific expression is due to transcriptional control [22]. It is also bound to the granule in many other organisms such as pea, potato, wheat, and *Chlamydomonas reinhardtii* [16, 23-25]. Edwards *et al.* [26] found that the C-terminal region of GBSSI in potato is different from the C-terminal region of other SSs. It also has specific properties pertaining to its affinity to ADPG and glucan substrates, activation by amylopectin, response to citrate, and thermo-sensitivity. These properties may be important in GBSSI's specificity to make amylose.

A second GBSS, GBSSII, was found in wheat and is encoded by completely different loci from GBSSI [27]; GBSSII is found primarily in the pericarp of wheat, while GBSSI is expressed in the endosperm. Vrinten and Nakamura's [27] research on GBSSII shows that genes related to GBSSII also occur in barley, rice, and maize. GBSSII is more sensitive to physiological conditions that do not affect GBSSI, and the presence of two genes may be an adaptive advantage over the single isoform of GBSS found in *Chlamydomonas* [27]. Further investigation of GBSSII is needed to determine its exact function in wheat and possibly other organisms.

Van de Wal *et al.* [28] has shown that amylose synthesis by GBSSI occurs by the elongation of available amylopectin chains within the amylopectin matrix in *Chlamydomonas reinhardtii*. A similar observation was made in potatoes, where GBSSI played a large role in amylopectin synthesis and granule morphology [29]. In pea embryos, evidence shows that GBSSI worked on the granule surface, while other enzymes are active in the soluble portion [30]. Because GBSSI is bound only in the granule, the method of synthesizing amylose from available amylopectin chains seems a likely theory. However, more current research by Denyer *et al.* [31] has shown that GBSSI synthesizes amylose by elongating malto-oligosaccharides (MOS) diffused into the granule matrix, instead of amylopectin. In fact, they show that amylopectin acts as an effector that stimulates the rate of elongation and increases the affinity of the enzyme for the MOS substrate. They also show that the presence of amylopectin causes GBSSI to change from a distributive enzyme to a processive enzyme. Denyer [32] proposed a model of amylose synthesis where GBSSI binds tightly to the amylopectin matrix where MOS and ADPG are also trapped. In this matrix, GBSS has a high affinity for MOS and elongates the substrate processively to make long glucan chains that are also trapped in the amylopectin matrix. Dauvillée *et al.* [33] show that GBSSI requires a primer and amylopectin matrix to be activated in *Chlamydomonas,* which supports Denyer's theory.

Starch synthase I

SSI is involved in amylopectin biosynthesis and to this date no mutants lacking SSI activity have been reported except in *Arabidopsis thaliana* where SSI was found to be crucial for the synthesis of normal amylopectin in leaves [34]. However, the properties of the SSI enzyme have been thoroughly explored through successful purification. SSI can catalyze "unprimed" glucan synthesis in the presence of high concentrations of citrate where other SS enzymes cannot [35]. SSI's contribution to amylopectin synthesis in maize is approximately 60-70% of the total soluble SS activity [36]. In rice endosperm, SSI accounts for the majority of total SS activity [37]. It has been associated with the 76 kDa protein found in the

soluble and granule-associated fractions of maize and wheat starch [38, 39]. A single copy of SSI has been found in the maize genome and is located on chromosome 9 near the *wx* locus [40]. Maize SSI has an N-terminal extension of 93 amino acids whose function is thought to aid in regulation of binding to α-glucans but it not a part of the catalytic domain [41]. Evidence from rice and maize endosperm indicated that there is a reduction in SSI activity in branching enzyme IIb mutants, *ae*, suggesting an association between the two enzymes [37, 42].

Commuri and Keeling [43] found that SSI is primarily involved in the extension of the shortest glucan chains in maize, DP <10, until it becomes entrapped in the granule. Similar evidence was found in *Chlamydomonas* where each SS appeared to be responsible for the synthesis of distinct size classes of glucans in all starch fractions [44]. Rice containing a retro-transposon inserted into the SSI gene resulted in a mutant with amylopectin depleted in chains of DP 8-12 and enriched chains of DP 6-7 indicating that rice SSI functions in the synthesis of chains with DP 8-12 from chains of DP 6-7 [45]. Commuri and Keeling [43] theorize that after SSI extends the short chains and that another SS, like SSIIa, takes over the chain elongation from there. According to this theory each SS is responsible for the extension of a certain glucan chain lengths in amylopectin synthesis in order to make the final product. This would explain the need for so many isoforms of SSs for amylopectin synthesis.

Starch synthase IIa and IIb

SSIIa and SSIIb are two SSs that are very similar in sequence but possess different enzymatic properties as well as dissimilar tissue locations. In barley and pea SSII is associated with the granule and the soluble fractions [14, 30]. Similar results are found in wheat where SSII is associated with granule and the soluble fractions during early development of the endosperm but is strictly found only in the granule fraction at mid to late development of the endosperm [46]. The accepted full length of maize SSIIa is 669 amino acids and SSIIb is 637 amino acids [47]. Harn *et al.* [19] deduced that clones of each shared

58.1% amino acid sequence identity and whose genes were found located on chromosome 6 near the *su2* locus. SSIIa is a major part of the total SS activity in pea and *Chlamydomonas* where in most cereals it only plays a minor part [48-52].

Expression studies showed maize SSIIa was located predominantly in the endosperm and a reduced amount in embryo and leaf [19]. Interestingly, maize SSIIb had strong transcript levels in leaf and weaker in endosperm and embryo [19]. Another difference between these two maize SSs is their preference of substrate. SSIIa demonstrated a higher V_{max} for amylopectin versus glycogen where SSIIb showed a higher V_{max} for glycogen versus amylopectin [47]. SSIIb displayed a much higher specific activity than SSIIa, but also had an approximate two-fold decrease affinity for either glucan primer [47].

Recently, Zhang *et al.* [53] discovered that *su2* codes for SSIIa in maize whose mutation suggests that SSIIa is required for the production of glucan chains between DP12 to DP25. The reduction of the activity of SSII in potato resulted in amylopectin chains with an enrichment of chains DP 8-12 and a decrease in chains DP 15-25 [54]. Similar data was found in rice, wheat, barely, pea, and *Chlamydomonas* where the lack of SSIIa activity resulted in an enrichment of shorter glucan chains and a decrease in intermediate glucan chains [48-52]. This supports the theory that each isoform of SS is responsible for the extension of certain glucan chain lengths in amylopectin synthesis.

Starch synthase III

SSIII is the last known SS in maize which is left to discuss and it is coded by the *du1* locus and has a predicted size of 188 kDa [36]. This maize SS is much larger compared to the previously discussed SSs due to a long N-terminal extension whose function is yet to be determined. Potato SSIII also has a long N-terminal extension and its proposed function is that it is involved in the interaction of starch polymers and does not have any interaction with catalytic activity [55]. Rice has duplicate forms of SSIII (SSIII-1 and SSIII-2) where SSIII-1 is primarily expressed in the endosperm and SSIII-2 is mainly expressed in the leaves [11]. SSIII is the major SS in potato tubers and accounts for approximately 80% of the total SS

activity in the soluble fraction of starch [55]. Purified SSIII from maize showed comparable V_{max} values for amylopectin and glycogen and even showed some activity using amylose as a glucan primer [56]. This evidence indicates that SSIII functions in the elongation of relatively longer glucan chains in amylopectin synthesis which again supports the theory that each SS isoform is responsible for the elongation of a certain number of glucose units.

Other enzymes involved in starch synthesis

Branching enzyme

There are two classes of branching enzyme (BE) found in maize endosperm including BEI, BEIIa, and BEIIb [4]. BE's is an amylo-(1, 4), (1, 6)-transglycosylase that cleaves α-1, 4 linkages and then joins the reducing end of the detached fragment to C-6 hydroxyls [6]. BEII has been found associated with the granule and soluble fractions in maize starch [57]. The two classes differ in the lengths of the chains transferred *in vitro*, BEII transfers shorter chains than BEI [4]. The action of each of the maize BE isoforms on amylopectin is different where BEI isoforms had the highest activity in branching amylose and much lower activity in branching amylopectin and BEII isoforms had the highest activity in branching amylopectin and lower activity in branching amylose [58]. These enzymes are required for the synthesis of amylopectin due to its role in creating amylopectin's α-1, 6 linkages and play a significant part in creating the granule structure.

Debranching enzyme

Two debranching enzyme (DBE) families exist in plants, isoamylase-type and pullulanase-type, and are responsible for the hydrolysis of α-1, 6 linkages [59]. The *sugary1* (*su1*) gene in maize was found to code for an isoamylase-type DBE and mutation in the *su1* gene resulted in an increase of sucrose concentration, decreased concentration of amylopectin, and an accumulation of highly branching glucopolysaccharide phytoglycogen [60]. *Zpu1* in maize codes for a pullulanase-type DBE and a null mutation of this gene

results in endosperm accumulating branched malto-oligosaccharides (MOS) and is deficient in linear MOS [59]. The role of DBE's in starch biosynthesis is not exactly clear but according to the "glucan-trimming" model, glucan trimming is required for amylopectin aggregation into an insoluble granular structure [3, 61]. This evidence indicates that the final branched product of amylopectin in starch is created through the combined efforts of BE's and DBE's.

Starch phosphorylase

Starch phosphorylase (SP) "catalyses the reversible transfer of glycosyl units from glucose-1-phosphate to the non-reducing end of α-1, 4-linked glucan chains and may be driven in either a synthetic or a degradative direction by the relative concentrations of the soluble substrates" [2]. Plastidial SP is called Pho1 and in maize has a higher affinity for the substrate amylopectin than glycogen [62]. Plastidial SP in maize is a 112 kDa protein found in the stroma of the amyloplast [62]. Mu *et al.* has shown that SP in maize prefers the phosphorolytic reaction versus a synthetic reaction when MOS were used as substrates [63]. Takaha *et al.* [64] suggest that SP may work together with disproportionating enzymes (D-enzymes) for starch synthesis using the SP phosphorolytic reaction in the "glucan-trimming" model [3, 61]. According to this model, pre-amylopectin molecules are trimmed by starch DBE releasing short chain MOS which are converted into long chain glucan molecules by the action of the D-enzyme. Then the longer chain glucans are exploited by SP via the phosphorolytic reaction to generate glucose-1-phosphate which can be used by AGP [64].

Disproportionating enzyme

The D-enzyme is a 4-α-glucanotransferase which transfers two glucosyl units from malto-triose onto a longer glucan chain [2]. Its primary function is thought to be involved in starch degradation but may also play a role in starch biosynthesis using SP as mentioned earlier [64]. Colleoni *et al.* have found that D-enzyme transfers glucans from among the outer chains of amylopectin and from MOS into the outer chains of amylopectin or glycogen

indicating an involvement of D-enzyme in starch synthesis [65]. Further research on D-enzyme is necessary in order completely understand its exact role in starch synthesis.

Conclusion

Starch is a semi-crystalline structure made up of amylose and amylopectin. It is synthesized in higher plants in order to store large amounts of glucose without osmotic consequences. The synthesis of starch is a complex reaction using an array of enzymes including (1) SSs which elongate glucan chains via α-1, 4 linkages; (2) BE's that create the α-1, 6 linkages in amylopectin; (3) DBE's which hydrolyze α-1, 6 linkages; (4) SP that transfer glycosyl units from glucose-1-phosphate to the non-reducing end of α-1, 4-linked glucan chains; and (5) D-enzyme which is a 4-α-glucanotransferase. The functional relationship between all of these enzymes is not completely understood. However, recent evidence indicates that some may form complexes with each other for regulation and/or activation purposes [2].

There is a lot of information known about the enzymes involved in starch synthesis; however exactly how semi-crystalline starch granules are formed is still unknown. Further understanding of all the enzymes involved in starch synthesis is needed to give clues to this mystery and may be easier to study once the entire maize genome is sequenced. As soon as starch synthesis is understood, starch may be manipulated for industrial applications. For example, if the chain distribution in amylopectin is altered it may result in a novel starch type whose application has already shown interest in industry. Research on SSs appear to be the best way to alter chain length distribution due to the evidence of each SS having a specific role in the elongation of various chains in the synthesis of amylopectin [43]. The manipulation of SSs may also reveal the exact function of each in starch synthesis. Once starch synthesis is understood improvements can be made to add more value to starch-storing crops.

References

1. Institute, I.S., *Starch Application Table*. 1999: Science Park Aarhus, Denmark.
2. Tetlow, I.J., M.K. Morell, and M.J. Emes, *Recent developments in understanding the regulation of starch metabolism in higher plants.* J Exp Bot, 2004. **55**(406): p. 2131-45.
3. Myers, A.M., et al., *Recent progress toward understanding biosynthesis of the amylopectin crystal.* Plant Physiol, 2000. **122**(4): p. 989-97.
4. James, M.G., K. Denyer, and A.M. Myers, *Starch synthesis in the cereal endosperm.* Curr Opin Plant Biol, 2003. **6**(3): p. 215-22.
5. Laughlin, M.J., S.E. Chantler, and T.W. Okita, *N- and C-terminal peptide sequences are essential for enzyme assembly, allosteric, and/or catalytic properties of ADP-glucose pyrophosphorylase.* Plant J, 1998. **14**(2): p. 159-68.
6. Buchanan, B.B., Gruissem, W., and Jones, R.L, *Biochemistry & Molecular Biology of Plants*. 2000, Rockville, MD: American Society of Plant Physiologists. 1367.
7. Ballicora, M.A., et al., *ADP-Glucose pyrophosphorylase from potato tubers. Site-directed mutagenesis studies of the regulatory sites.* Plant Physiol, 1998. **118**(1): p. 265-74.
8. Beckles, D.M., A.M. Smith, and T. ap Rees, *A cytosolic ADP-glucose pyrophosphorylase is a feature of graminaceous endosperms, but not of other starch-storing organs.* Plant Physiol, 2001. **125**(2): p. 818-27.
9. Denyer, K., et al., *The major form of ADP-glucose pyrophosphorylase in maize endosperm is extra-plastidial.* Plant Physiol, 1996. **112**(2): p. 779-85.
10. Shannon, J.C., et al., *Brittle-1, an adenylate translocator, facilitates transfer of extraplastidial synthesized ADP--glucose into amyloplasts of maize endosperms.* Plant Physiol, 1998. **117**(4): p. 1235-52.
11. Dian, W., H. Jiang, and P. Wu, *Evolution and expression analysis of starch synthase III and IV in rice.* J Exp Bot, 2005. **56**(412): p. 623-32.

12. Nelson, O.E. and H.W. Rines, *The enzymatic deficiency in the waxy mutant of maize.* Biochem Biophys Res Commun, 1962. **9**: p. 297-300.

13. Shure, M., S. Wessler, and N. Fedoroff, *Molecular identification and isolation of the Waxy locus in maize.* Cell, 1983. **35**(1): p. 225-33.

14. Li, Z., et al., *The structural organisation of the gene encoding class II starch synthase of wheat and barley and the evolution of the genes encoding starch synthases in plants.* Funct Integr Genomics, 2003. **3**(1-2): p. 76-85.

15. MacGregor, E.A., *Possible structure and active site residues of starch, glycogen, and sucrose synthases.* J Protein Chem, 2002. **21**(4): p. 297-306.

16. Ainsworth, C., J. Clark, and J. Balsdon, *Expression, organisation and structure of the genes encoding the waxy protein (granule-bound starch synthase) in wheat.* Plant Mol Biol, 1993. **22**(1): p. 67-82.

17. Mason-Gamer, R.J., C.F. Weil, and E.A. Kellogg, *Granule-bound starch synthase: structure, function, and phylogenetic utility.* Mol Biol Evol, 1998. **15**(12): p. 1658-73.

18. Gao, Z., et al., *Involvement of lysine-193 of the conserved "K-T-G-G" motif in the catalysis of maize starch synthase IIa.* Arch Biochem Biophys, 2004. **427**(1): p. 1-7.

19. Harn, C., et al., *Isolation and characterization of the zSSIIa and zSSIIb starch synthase cDNA clones from maize endosperm.* Plant Mol Biol, 1998. **37**(4): p. 639-49.

20. Commuri, P.K., Peter L.; Ramirez, Nona; McKean, Angela; Gao, Zhong; Guan, Hanping, *Glucan Chain Length Domains*, in *United States Patent Application.* 2004.

21. Kelley, L.A., R.M. MacCallum, and M.J. Sternberg, *Enhanced genome annotation using structural profiles in the program 3D-PSSM.* J Mol Biol, 2000. **299**(2): p. 499-520.

22. Klosgen, R.B., et al., *Molecular analysis of the waxy locus of Zea mays.* Mol Gen Genet, 1986. **203**: p. 237-244.

23. Smith, A.M., *Evidence that the "waxy" protein of pea (Pisum sativum L.) is not the major starch-granule-bound starch synthase.* Planta, 1990. **182**: p. 599-604.

24. Visser, R.G., et al., *Inhibition of the expression of the gene for granule-bound starch synthase in potato by antisense constructs.* Mol Gen Genet, 1991. **225**(2): p. 289-96.

25. Delrue, B., et al., *Waxy Chlamydomonas reinhardtii: monocellular algal mutants defective in amylose biosynthesis and granule-bound starch synthase activity accumulate a structurally modified amylopectin.* Journal of Bacteriology, 1992. **174**(1): p. 3612-3620.

26. Edwards, A., et al., *Specificity of starch synthase isoforms from potato.* Eur J Biochem, 1999. **266**(3): p. 724-36.

27. Vrinten, P.L. and T. Nakamura, *Wheat granule-bound starch synthase I and II are encoded by separate genes that are expressed in different tissues.* Plant Physiol, 2000. **122**(1): p. 255-64.

28. van de Wal, M., et al., *Amylose is synthesized in vitro by extension of and cleavage from amylopectin.* J Biol Chem, 1998. **273**(35): p. 22232-40.

29. Fulton, D.C., et al., *Role of granule-bound starch synthase in determination of amylopectin structure and starch granule morphology in potato.* J Biol Chem, 2002. **277**(13): p. 10834-41.

30. Denyer, K., et al., *Soluble isoforms of starch synthase and starch-branching enzyme also occur within starch granules in developing pea embryos.* Plant J, 1993. **4**(1): p. 191-8.

31. Denyer, K., et al., *Interaction with amylopectin influences the ability of granule-bound starch synthase I to elongate malto-oligosaccharides.* Biochem J, 1999. **342 Pt 3**: p. 647-53.

32. Denyer, K., *The control of amylose synthesis.* Journal of Plant Physiology, 2001. **158**: p. 479-487.

33. Dauvillee, D., et al., *Novel, starch-like polysaccharides are synthesized by an unbound form of granule-bound starch synthase in glycogen-accumulating mutants of Chlamydomonas reinhardtii.* Plant Physiol, 1999. **119**(1): p. 321-30.

34. D'Hulst, C. (2005). Université des Sciences et Technologies de Lille, Villeneuve d'Ascq, France.

35. Boyer, C.D. and J. Preiss, *Properties of citrate-stimulated starch synthesis catalyzed by starch synthase I of developing maize kernels.* Plant Physiol, 1979. **64**: p. 1039-1042.

36. Cao, H., et al., *Identification of the soluble starch synthase activities of maize endosperm.* Plant Physiol, 1999. **120**(1): p. 205-16.

37. Nishi, A., et al., *Biochemical and genetic analysis of the effects of amylose-extender mutation in rice endosperm.* Plant Physiol, 2001. **127**(2): p. 459-72.

38. Mu, C., et al., *Association of a 76 kDa polypeptide with soluble starch synthase I activity in maize (cB73) endosperm.* The Plant Journal, 1994. **6**(2): p. 151-159.

39. Li, Z., et al., *Cloning and characterization of a gene encoding wheat starch synthase I.* Theor Appl Genet, 1999. **98**: p. 1208-1216.

40. Knight, M.E., et al., *Molecular cloning of starch synthase I from maize (W64) endosperm and expression in Escherichia coli.* Plant J, 1998. **14**(5): p. 613-22.

41. Imparl-Radosevich, J.M., et al., *Purification and characterization of maize starch synthase I and its truncated forms.* Arch Biochem Biophys, 1998. **353**(1): p. 64-72.

42. Boyer, C.D. and J. Preiss, *Multiple forms of starch branching enzyme of maize: evidence for independent genetic control.* Biochem Biophys Res Commun, 1978. **80**: p. 169-175.

43. Commuri, P.D. and P.L. Keeling, *Chain-length specificities of maize starch synthase I enzyme: studies of glucan affinity and catalytic properties.* Plant J, 2001. **25**(5): p. 475-86.

44. Maddelein, M.L., et al., *Toward an understanding of the biogenesis of the starch granule. Determination of granule-bound and soluble starch synthase functions in amylopectin synthesis.* J Biol Chem, 1994. **269**(40): p. 25150-7.

45. Nakamura, Y., *Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: rice endosperm as a model tissue.* Plant Cell Physiol, 2002. **43**(7): p. 718-25.

46. Li, Z., et al., *The localization and expression of the class II starch synthases of wheat.* Plant Physiol, 1999. **120**(4): p. 1147-56.

47. Imparl-Radosevich, J.M., et al., *Analysis of purified maize starch synthases IIa and IIb: SS isoforms can be distinguished based on their kinetic properties.* Arch Biochem Biophys, 1999. **362**(1): p. 131-8.

48. Yamamori, M., et al., *Genetic elimination of a starch granule protein, SGP-1, of wheat generates an altered starch with apparent high amylose.* Theor Appl Genet, 2000. **101**: p. 21-29.

49. Umemoto, T., et al., *Mapping of a gene responsible for the difference in amylopectin structure between japonica-type and indica-type rice varieties.* Theor Appl Genet, 2002. **104**(1): p. 1-8.

50. Morell, M.K., et al., *Barley sex6 mutants lack starch synthase IIa activity and contain a starch with novel properties.* Plant J, 2003. **34**(2): p. 173-85.

51. Craig, J., et al., *Mutations in the gene encoding starch synthase II profoundly alter amylopectin structure in pea embryos.* Plant Cell, 1998. **10**(3): p. 413-26.

52. Fontaine, T., et al., *Toward an understanding of the biogenesis of the starch granule. Evidence that Chlamydomonas soluble starch synthase II controls the synthesis of intermediate size glucans of amylopectin.* J Biol Chem, 1993. **268**(22): p. 16223-30.

53. Zhang, X., et al., *Molecular characterization demonstrates that the Zea mays gene sugary2 codes for the starch synthase isoform SSIIa.* Plant Mol Biol, 2004. **54**(6): p. 865-79.

54. Edwards, A., et al., *A combined reduction in activity of starch synthases II and III of potato has novel effects on the starch of tubers.* The Plant Journal, 1999. **17**(3): p. 251-261.

55. Marshall, J., et al., *Identification of the major starch synthase in the soluble fraction of potato tubers.* Plant Cell, 1996. **8**(7): p. 1121-35.

56. Cao, H., M.G. James, and A.M. Myers, *Purification and characterization of soluble starch synthases from maize endosperm.* Arch Biochem Biophys, 2000. **373**(1): p. 135-46.

57. Mu-Forster, C., et al., *Physical association of starch biosynthetic enzymes with starch granules of maize endosperm. Granule-associated forms of starch synthase I and starch branching enzyme II.* Plant Physiol, 1996. **111**(3): p. 821-9.

58. Guan, H.P. and J. Preiss, *Differentiation of the Properties of the Branching Isozymes from Maize (Zea mays).* Plant Physiol, 1993. **102**(4): p. 1269-1273.

59. Dinges, J.R., et al., *Mutational analysis of the pullulanase-type debranching enzyme of maize indicates multiple functions in starch metabolism.* Plant Cell, 2003. **15**(3): p. 666-80.

60. James, M.G., D.S. Robertson, and A.M. Myers, *Characterization of the maize gene sugary1, a determinant of starch composition in kernels.* Plant Cell, 1995. **7**(4): p. 417-29.

61. Ball, S., et al., *From glycogen to amylopectin: a model for the biogenesis of the plant starch granule.* Cell, 1996. **86**(3): p. 349-52.

62. Yu, Y., et al., *Identification of the maize amyloplast stromal 112-kD protein as a plastidic starch phosphorylase.* Plant Physiol, 2001. **125**(1): p. 351-9.

63. Mu, H.H., et al., *Purification and characterization of the maize amyloplast stromal 112-kDa starch phosphorylase.* Arch Biochem Biophys, 2001. **388**(1): p. 155-64.

64. Takaha, T., et al., *Disproportionating enzyme (4-alpha-glucanotransferase; EC 2.4.1.25) of potato. Purification, molecular cloning, and potential role in starch metabolism.* J Biol Chem, 1993. **268**(2): p. 1391-6.

65. Colleoni, C., et al., *Biochemical characterization of the chlamydomonas reinhardtii alpha-1,4 glucanotransferase supports a direct function in amylopectin biosynthesis.* Plant Physiol, 1999. **120**(4): p. 1005-14.

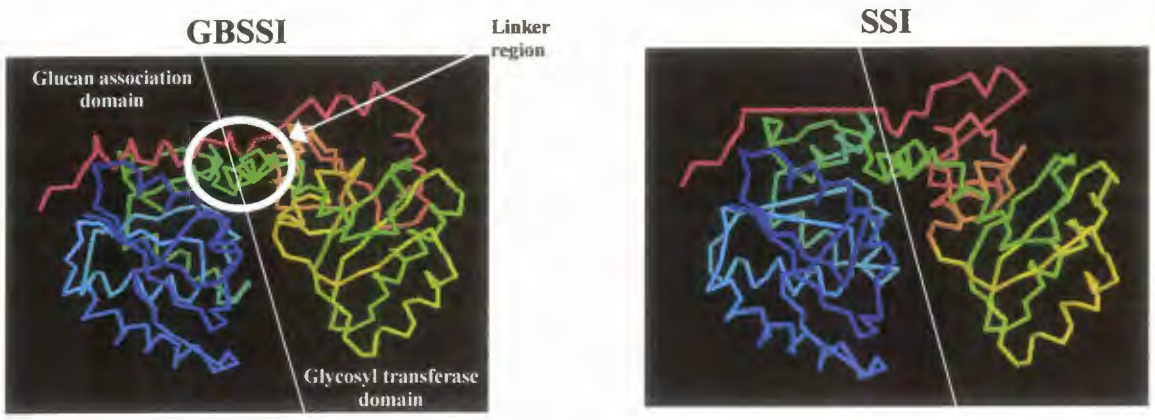


Figure 1: 3D-PSSM predicted structural models of GBSSI and SSI [21].

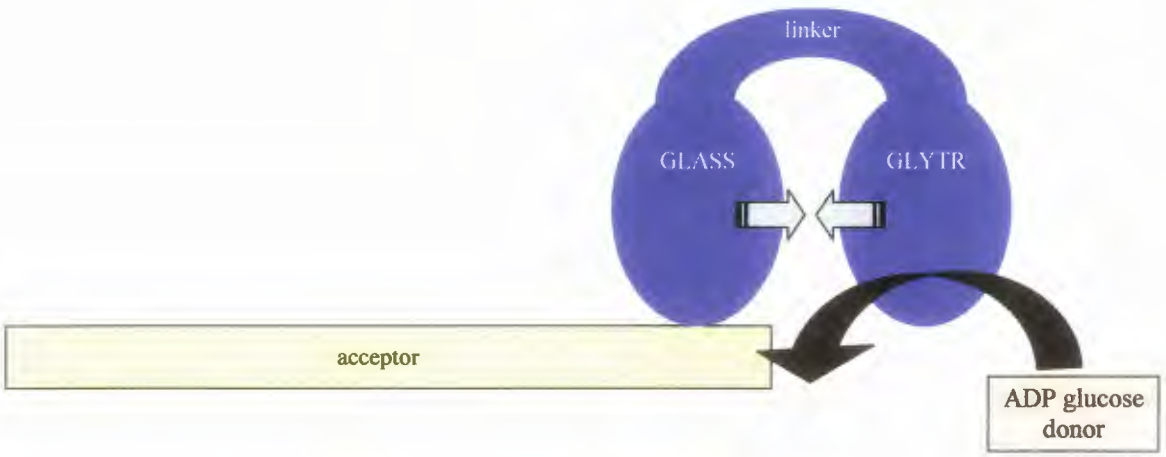


Figure 3: Schematic representation of the GLASS/GLYTR hypothesis.

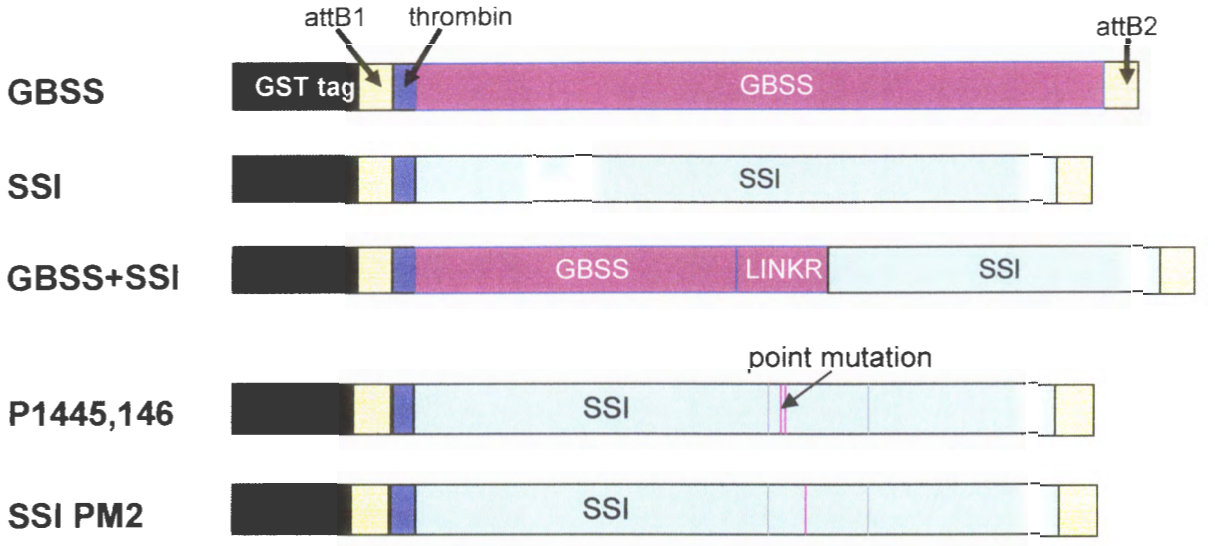


Figure 3: Visual representation of the five plasmids` construction.

Thesis Organization

The research reported in this thesis was written by me, with editing from my major professors Dr. Alan Myers and Dr. Martha James, and all experiments were performed by me. It contains information from my main research project and from two other "side" projects. The main research chapter is Chapter 2 and is a description of the investigation of the structure and function of maize GBSS and SSI. The references and figures are found at the end of the text. Note that a couple of the figures are repeated from Chapter 1 to make it easier to read and refer to the figures without having to go back to Chapter 1.

There are three appendices following the main text. Appendix A contains the plasmid sequences of the vectors used to perform the research in Chapter 2. Appendix B contains additional research on SSs, namely *su2*'s impact on the proteins entrapped in the starch granule. Appendix C is also additional research performed during lab rotations in Dr. Mark Westgate's lab and investigates maize dehydrin levels in varying degrees of drought-stressed plants. The research on dehydrins was an extension of the work performed by Octavio Caviglia.

List of Abbreviations

ADPG	adenosine 5' diphosphate glucose
AGP	adenosine 5' diphosphate glucose pyrophosphorylase
BE	branching enzyme
BSA	bovine serum albumin
DAP	days after pollination
DBE	debranching enzyme
D-enzyme	disproportionating enzyme
DP	degree of polymerization
E. coli	*Escherichia coli*
GBSS	granule bound starch synthase
GLASS	glucan association

GLYTR	glycosyl transferase
GST	glutathione-S-transferase
Hsp	heat shock protein
MOS	malto-oligosaccharide
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PM	point mutation
PMSF	phenyl methyl sulfonyl fluoride
Rubisco	ribulose bisphosphate carboxylase
SDS	sodium dodecyl sulfate
SP	starch phosphorylase
SSI	starch synthase I
SSIIa	starch synthase IIa
SSIIb	starch synthase IIb
SSIII	starch synthase III
Wx	*Waxy*

CHAPTER 2. FUNCTIONAL ANALYSIS OF MAIZE GBSSI AND SSI: CHARACTERIZATION OF RECOMBINANT CHIMERIC PROTEINS

Introduction

Starch is a commodity used by various manufacturers for a variety of purposes. For example, starch is the primary carbohydrate component in human diets and is also used in livestock feed. Some non-food uses include starch as an ingredient in pharmaceuticals, film coating, packing material, cosmetics, and other important industrial applications [1]. Starch is synthesized in plants for the purpose of storage of carbohydrate reserves. However, the exact mechanism of starch synthesis in plants is not fully understood. This study focuses on understanding the molecular mechanisms of maize starch biosynthesis because the primary source of starch in the United States is harvested from corn endosperm.

Amylose and amylopectin are glucan polymers that comprise the starch granule. Amylose is a predominantly linear polymer of α-1, 4 linked glucose units. Amylopectin has the most abundant polymer and contains α-1, 4 linked glucose chains of varying length due to approximately 5% branching in which the linkages are α-1, 6 linkages[2]. Amylopectin is a highly organized structure in which the linear glucan chains are clustered and alternate with highly branched regions. This organization of α-1, 4 and α-1, 6 linkages is responsible for the semi-crystalline structure of starch, which enables the dense packing of a large number of glucose units. In order to better understand how semi-crystalline starch is formed, this study focuses on the specific enzymatic functions of two starch synthases (SS's) required for synthesizing the linear amylose and amylopectin chains.

SS's are enzymes that catalyze the elongation of glucose chains via the introduction of α-1, 4 linkages during starch biosynthesis. There are at least five known SS's in maize, including granule bound starch synthase (GBSS), starch synthase I (SSI), starch synthase IIa (SSIIa), starch synthase IIb (SSIIb), and starch synthase III (SSIII). GBSS functions mostly in the elongation of amylose [3, 4] while the remaining SS's are involved in amylopectin

biosynthesis [5]. The structures of all SS's, along with glycogen and sucrose synthases, are predicted to contain of two essential activity domains required for starch binding and catalysis [6]. A linker region, which is a stretch of amino acid residues between these two domains, allows for the flexibility needed for the enzyme to attain the opened and closed formations [7]. When the maize SS primary amino acid sequences were threaded onto a protein threading program called 3D-PSSM [8] (see Figure 1), it was found that these proteins are highly comparable with the highest confidence to UDP-N- acetylglucosamine 2-epimerase which belong to a family called glycosyltransferase group 1, or Pfam 00534 [7].

Commuri *et al.* [6] identified the two activity domains of maize SS's as the glucan association domain (GLASS) and glycosyl transferase domain (GLYTR). These authors also proposed that the function of the GLASS domain possesses binding properties (binding the amylopectin molecule in a growing starch granule), whereas the GLYTR domain may be involved in interaction with the glucan during the catalytic process of α-1, 4 bond formation during glucan chain elongation (see Figure 2).

This present study aims to investigate how the interaction between the two domains effects the degree of functionality by generating and expressing a chimeric SS in *E. coli* cells. The chimeric SS would consist of a fusion of the GLASS domain of GBSS with the GLYTR domain of SSI. SSI was selected for recombination because it contributes approximately 60-70% of the total SS activity in maize [9] and is found in both the granule and soluble fraction of plastid [10]. Commuri and Keeling [11] found that maize SSI is primarily involved in the extension of the shortest glucan chains in maize, DP<10, until it becomes entrapped in the granule. GBSS was chosen in that it quite different from the other SS's because it is found only in the granule fraction of maize starch [12]. The primary function of maize GBSS is the elongation of amylose and does so by the addition of glucans to malto-oligosaccharides (MOS) using amylopectin as an effector to stimulate the rate of elongation along with increasing the affinity of the enzyme to the MOS substrate [13]. The two enzymes also have varying enzymatic properties. The primary binding substrate of GBSSI is amylose where SSI mainly binds glycogen based off the previously investigated K_d values (see Table 1) [6].

The combination of the GBSSI GLASS and SSI GLYTR domains could result in an enzyme with a high affinity for amylose (GBSSI) and the ability to extend glucan chains of DP<10 (SSI).

Chimeric SS's have been previously investigated in potatoes by Edwards *et al.* [14]. They discovered features specific to the GBSSI enzyme including MOS elongation and amylopectin stimulation. However, they did not know about the GLASS and GLYTR domains at that time and constructed truncated and chimeric versions of the potato GBSSI and SSII isoforms through expression in *E. coli* based solely on amino acid composition. The kinetics, themosensitivity, specific activity, and product specificity of their enzymes were evaluated with the hopes of identifying regions that confer those particular properties [14]. Only a few of their constructs had measurable activity, but those that did revealed some clues about the function of those specific regions of those enzymes. If chimeric SS's can be recombined while retaining activity, then enzymes could be expressed in maize with the potential of creating starch with altered chain length distributions with industrial application becomes possible. In addition to starch with novel properties, the exact function of SS domains may be revealed due to the now identified GLASS and GLYTR domains. The hypothesis of this research is that GLASS and GLYTR interaction affects the enzymes' degree of functionality (see Figure 2).

Rearrangement of SS's in the linker region may cause the enzyme to be inactive; therefore, testing the susceptibility of the linker region to alteration will also be examined by introducing point mutations in that region of a wild type SSI enzyme. The theory of how GLASS and GLYTR interact is based on the fact that there is flexible linker region between the two. If this region were to become more or less flexible it may have an effect on the enzyme's ability to act on its substrate. The hypothesis is that minor rearrangement in the linker region of SSI does not cause a loss of activity of the enzyme because the extent of rearrangement is small.

This report investigates the interaction of the GLASS and GLYTR domains effects on an enzyme's degree of functionality to act on a substrate by combining the GLASS and

linker domains of GBSS with the GLYTR domain of SSI. The results indicate that these two domains most likely interact in order for the enzyme to function. The susceptibility of the linker region is also investigated to determine the consequences of recombination in this region. The results indicate that this domain can withstand minor alteration without the loss of activity.

Materials and Methods

Plasmid construction

A total of five plasmids were constructed using Invitrogen's Gateway™ Cloning Technology (see Figure 3) [15]. cDNA sequence of each SS was PCR amplified using Platinum Pfx (Invitrogen) according to manufacturer's instructions using the supplied buffers (final concentration of $MgCl_2$ was 2 mM) from plasmid clones of the cDNA sequence, pEXSwx2 (full length GBSS) and pEXS1d (full length of SSI catalytic domain), which were generous gifts from Dr. Peter Keeling and Dr. Hanping Guan from BASF-Plant Sciences, Ames, IA. SS enzymes were PCR amplified using sequence-specific primers containing Gateway's specified recombination sites for bacteriophage lambda (λ) in *E. coli* cells which are required for cDNA insertion [15]. A thrombin cleavage site was included on the N-terminal primer to enable the eventual cleavage of the GST-tag from the purified protein. The PCR fragment was purified on a 1% agarose gel with ethidium bromide DNA was purified using Qiagen's gel extraction kit by following the provided protocol and eluting with 50 µl sterile water. *In vitro* recombination was performed using the recombination sites *att*B and *att*P (BP) for bacteriophage λ in *E. coli* cells and was incubated overnight at room temperature [15]. Reaction contents contained 4 µl BP reaction buffer, 4 µl purified PCR product, 2 µl pDONR 221, 6 µl 1x TE, and 4 µl BP clonase. To confirm the identity of the insertion of the cDNA into the DONR vector, plasmids were sequenced by the Iowa State University DNA Sequencing Facility. Sequencing confirmed an exact match to the template DNA and would code for an in-frame predicted protein. Each insert was then transferred from the DONR plasmid via an LR (*att*L and *att*R recombination sites) recombination

reaction (3 µl of 1/10 dilution of entry/DONR vector, 4 µl pDEST15, 4 µl LR reaction buffer, 5 µl 1 x TE, and 4 µl LR recombinase) to an Invitrogen expression plasmid, pDEST15, designed for the expression of an N-terminal glutathione-S-transferase (GST) tag fusion protein in *E. coli* cells [15] in an overnight incubation at room temperature.

Expression and Purification

Each construct was expressed in the Tuner *E. coli* strain from Novagen by transforming the plasmid into the competent strain. Tuner strains are *lacYZ* deletion mutants of BL21 and enable adjustable levels of protein expression throughout all the cells in a culture. Single colonies were used to inoculate 20 ml of LB containing 50 µg/ml carbenicillin and were incubated overnight at 37°C while shaking (210 rev/min). The entire 20 ml of overnight culture was then used to subculture into 1 L of LB containing 50 µg/ml carbenicillin and was grown at 37°C while shaking (210 rev/min) to an OD_{595} of approximately 0.4. Protein expression of the T7 promoter was initiated by the addition of 1 mM IPTG and then was placed in a 25°C incubator to shake (210 rev/min) for 2.5 hours. Cells were harvested by centrifugation at 5000*g* and pellets were gently resuspended in 50 ml of cold PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) containing 1 mM PMSF. Cells were lysed in ice using an Artek sonic dismembrator model 150 with 4 x 1 min bursts with 30 sec rest in between. Sonicate was then centrifuged at 55,500*g* and the supernatant containing the protein of interest was collected and placed on ice.

Affinity chromatography was used to isolate the GST-fusion protein from the *E. coli* extract. 3 ml of 50% glutathione sepharose™ 4B from Amersham Biosciences was added to the collected supernatant and then rocked gently overnight at 4°C. Bead/supernatant mixture was then poured into a 10 ml disposable column and the beads were allowed to settle (Note: all purification steps were performed at 4°C unless otherwise noted). Then the column was opened and the supernatant was allowed to flow-through. Afterwards, it was washed three times with 15 ml of cold PBS. GST-fusion proteins were at room temperature by the addition of 1.5 ml of fresh, room temperature glutathione elution buffer (0.154 g reduced

glutathione dissolved in 50 ml of 50 mM Tris-HCl, pH 8.0) to the closed column for 10 min. Then the elution buffer was allowed to flow through the column and was collected and placed on ice after the entire volume of the buffer had passed through. This elution step was repeated another two times to ensure that all of the GST fusion protein eluted off of the column. Column fractions were pooled and concentrated using Amicon 500 µl concentrators (YM 30,000).

The protein concentration of each fraction was determined using a Bradford Assay with bovine serum albumin (BSA) for a standard curve [16]. Next, 150 µl of each eluate was used to perform the thrombin cleavage reaction to remove the GST-tag from the protein of interest. Biotinylated thrombin from Novagen was added to the 150 µl of eluate based on protein concentration (1 U thrombin/1 mg protein) and was incubated at room temperature overnight. The cleaved GST-tag was removed by the addition of 66.5 µl of 50% glutathione sepharose™ 4B and allowing it to rock gently at room temperature for 30 min. Beads were centrifuged down at maximum speed in a microcentrifuge and the supernatant containing the protein of interest and biotinylated thrombin was collected. 25 µl of 50% streptavidin agarose beads from Novagen was added to the collected supernatant and was allowed to rock gently at room temperature for 30 min to remove the biotinylated thrombin. The streptavidin beads were removed by centrifugation at maximum speed in a microcentrifuge and the supernatant was collected which contained the purified protein of interest. The concentration of the purified proteins was determined again by using the Bradford Assay [16] with a BSA standard curve.

Immunodetection

Proteins were loaded in equivalent amounts onto 7.5% SDS-PAGE Bio-Rad pre-cast gels and run at 150 volts at room temperature in a Bio-Rad Mini-Protean II cell until the dye front ran off according to methods described by Laemmli [17]. After separation, proteins were electroblotted onto a nitrocellulose membrane in Towbins (25mM Tris-acetate (pH = 8.3), 192 mM glycine, and 20% methanol) at 325-350 mA while stirring for 45 min.

Membranes were blocked in 5% skim milk in Tris buffered saline with 0.1% Tween 20 (TBST: 10 mM Tris (pH = 7.5), 150 mM $NaCl_2$, and 0.1% Tween 20) shaking gently for 1 hour at room temperature. Blocking solution was discarded and then the membrane was incubated in 1° antibody shaking gently for 45 min at room temperature. 1° antibodies used in these experiments varied depending on the enzyme being analyzed and specific details can be found in the results section. Following 1° antibody incubation, membrane was washed three times in shaking gently in approximately 50 ml of TBST for 15 min at room temperature. Once the wash steps were completed, the membrane was incubated in 2° antibody (1:5000) (anti-rabbit IGg with horse radish peroxidase (HRP) conjugate from Santa Cruz Biotechnology) shaking gently for 30 min at room temperature. Then the three wash steps with TBST were repeated. Two more wash steps with TBS shaking gently for 10 min each at room temperature were performed to remove excess Tween 20 which can interfere with ECL development. For membrane detection, an ECL development kit from Amersham Biosciences was used. Once the reaction was complete, the membranes were exposed to X-ray film and developed using standard X-ray development solutions and procedures.

Amino acid sequencing

N-terminal amino acid sequencing by Edman degradation was performed by the Iowa State University Protein Sequencing and Synthesis facility on the purified samples of wild-type GBSS and SSI. Purified proteins were separated using 7.5% SDS-PAGE Bio-Rad pre-cast gels at 150 volts at room temperature in a Bio-Rad Mini-Protean II cell until the dye front ran off [17]. Following electrophoresis, the SDS-PAGE gel was washed at room temperature in transfer buffer (25 mM Tris-HCl (pH 8.2), 192 mM glycine, 5% (v/v) methanol) for 15-20 min. Gloves were worn at all times when handling the gels and membranes to prevent oils and secretions from the skin from contaminating the samples. Whatman 3MM filter paper and PVDF Membrane were cut to the exact size of the SDS-PAGE gel. The two filter papers were soaked in a small amount of transfer buffer, and the PVDF was prewet in 5-10 ml of 100% (v/v) methanol for 5 sec and then soaked in 500 ml of

dH_2O for 5 min to remove the methanol. PVDF was equilibrated in 500 ml of transfer buffer for 10-15 min prior to use in blotting. Proteins from SDS-PAGE gel were electroblotted onto the PVDF membrane in the transfer buffer at 70 V for 1.5 hours at room temperature while stirring.

Once the transfer was completed, the membrane was stained using approximately 20 ml of Coomassie stain (1.2 mM Coomassie brilliant blue R-250, 40% (v/v) methanol, and 17% (v/v) glacial acetic acid) shaking gently for 30 min at room temperature. Membrane was destained in destain I (40% (v/v) methanol and 10% (v/v) glacial acetic acid) gently shaking for 20 min at room temperature and then in destain II (5% (v/v) methanol and 10% (v/v) glacial acetic acid) gently shaking for 1 hour. Membrane was dried and submitted to the sequencing facility after the bands of interest were identified.

Zymogram analysis

A modified zymogram was used to visualize protein activity by adding 2.5 µg of purified protein to each lane using the methods described by Cao *et al.* [9]. Purified protein from *E. coli* expression was boiled for 10 min in SDS-PAGE loading buffer (65mM Tris-HCl (pH = 6.8), 2% SDS, 10% glycerol, and 5% 2-mercaptoethanol) allowed to cool and then centrifuged at 10,000 rpm at room temperature for 5 min. Electrophoresis was performed under denaturing conditions (25mM Tris-HCl (pH = 8.3), 192 mM glycine, 0.1% SDS, and 5 mM DTT) until the dye front ran off (approximately 2.5-3 h) at 80 V at 4°C in an 8% acrylamide gel containing 0.1% rabbit liver glycogen in a Bio-Rad Mini-Protean II cell. The gel was washed four times at room temperature in 40 mM Tris-HCl (pH = 7.0) and 5 mM DTT for 30 min each to remove SDS from the gel. Gel was sealed (Seal-O-Meal) in a small pouch containing reaction buffer (100 mM Bicine (pH 8.0), 0.5 M citrate, 25 mM postasium acetate, 0.5 mg/ml BSA, 5 mM ADPG, 5 mM 2-mercaptoethanol, and 20 mg/ml rabbit liver glycogen) and incubated at room temperature for 36-48 h while shaking gently. Enzyme activity was detected by the addition of iodine stain (0.2% iodine and 2% potassium iodide in 10 mM HCl) and photographed immediately. Unfortunately, this type of activity analysis

has not shown success with GBSS activity most likely due the lack of a semi-crystalline substrate to act upon.

Quantification of activity

Quantification of activity was performed using slightly altered methods defined by Denyer *et al.* [13] using the starch synthase activity assay. Assays contained 100 mM Bicine-NaOH (pH 8.0), 5 mM EDTA, 0.5 M sodium citrate, 25 mM potassium acetate, 10 mM DTT, 0.5 mg/ml BSA, 1 mM ADP [U-^{14}C]glucose, varying concentrations of either amylopectin or rabbit liver glycogen, and 0.5 μg/ml SS enzyme and were incubated at room temperature for 20 min. Reaction was terminated by the addition of 1 ml 75% (v/v) methanol/1% (v/v) KCl and was mixed by vortexing. Glucan was allowed to precipitate for 5 min and then centrifuged at 5000*g* for 5 min. Supernatant was carefully removed and discarded. Pellet was washed by dissolving in 200 μl dH_2O by vortexing. Glucan was precipitated again for 2 min at room temperature with the addition of 1 ml of the 75% (v/v) methanol/1% (v/v) KCl and vortexed. Solution was centrifuged at 5,000*g* for 1 min and the supernatant was carefully removed and discarded. Pellet was dissolved in 500 μl water by vortexing and transferred to a scintillation vial. After the addition of 10 ml of Scinverse, the samples were counted using a 1600 TR liquid scintillation analyzer from Packard. If the SS enzyme was active then an incorporation of the radiolabeled ADPG should be detected. Unfortunately, this type of activity analysis is very difficult and has not shown much success with GBSS activity.

Affinity electrophoresis

In order to determine the enzymes' affinities to rabbit liver glycogen was affinity electrophoresis was performed using methods defined by Commuri and Keeling [11]. Equal concentrations of purified proteins (approximately 2.5 μg/ml) were suspended in loading buffer (.025% (w/v) bromophenol blue and 3% (v/v) glycerol) and loaded onto native PAGE gels (6% acrylamide in 1.5 M Tris-HCl buffer (pH = 8.8)) containing various concentrations

of rabbit liver glycogen (0%, 0.5%, 1.0%, 1.5%, 2.0%, and 2.5%). Electrophoresis was carried out at 80V at room temperature in running buffer (25 mM Tris, 192 mM glycine, and 1 mM DTT) for exactly 2 h. The gels were stained with Coomassie stain for 30 min and destained with destain I for 20 min and destain II for 1 h. Migration distances of the protein bands were measured and recorded.

Results

Plasmid construction

To test the functionality of SS enzymes, five plasmids were created for the expression in *E. coli* of wild type GBSS and SSI enzymes, a chimeric enzyme comprising the GBSS GLASS and SSI GLYTR domains, and two enzymes containing point mutations in the SSI linker region. The cDNA of each was PCR amplified and then inserted into pDONR221 (Invitrogen) via recombination sites of the bacteriophage λ. Once the insertion was verified by DNA sequencing, the cDNA insert in the DONR vector was inserted into the *E. coli* expression vector pDEST15. Please refer to the following section for details on expression and purification.

Wild-type maize GBSSI cDNA was PCR amplified from a plasmid pEXSwx2 received from BASF-Plant Sciences, Ames, IA. PCR conditions were as follows: step 1 = 95°C for 10 min, step 2 = 95°C for 1 min, step 3 = 60°C for 1 min, step 4 = 72°C for 3 min, step 5 = repeat steps 2-4 35 times, and step 6 = 72°C for 5 min. The GBSSI-specific sequences in the PCR which flank the entire catalytic domain of the enzyme and include the GLASS and GLYTR domains (Genbank accession no. X03035) [6] are: 1) 5' G GGG ACA AGT TTG TAC AAA AAA GCA GGC TTA CTG GTG CCA CGC GGT AGT ATG AAC GTC GTC TTC GTC GGC 3' and 2) 5' GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TCA GGG CGC GGC CAC GTT CTC 3'. PCR product was inserted into pDONR221 for propagation and sequencing and then from pDONR211 to pDEST15. The

pDEST15+GBSS expression plasmid contains wild-type maize GBSSI cDNA which was expressed and purified from *E. coli* (see Figure 4).

Wild-type maize SSI was PCR amplified from plasmid pEXS1, also received from BASF-Plant Sciences. The SSI sequence within this plasmid codes for a polypeptide with an N-terminal truncation but which retains the catalytic activity of wild-type SSI [18]. PCR conditions were as follows: step 1 = 95°C for 10 min, step 2 = 95°C for 1 min, step 3 = 60°C for 1 min, step 4 = 72°C for 3 min, step 5 = repeat steps 2-4 35 times, and step 6 = 72°C for 5 min. The following primers flanking the entire catalytic SSI domain (including GLASS and GLYTR) were used (Genbank accession no. AF036891) [6]: 1) 5' G GGG ACA AGT TTG TAC AAA AAA GCA GGC TCC CTG GTG CCA CGC GGT AGT ATG AGC ATT GTC TTT GTA ACC 3' and 2) 5' GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TTA CAT GAC ATA GGG TCG 3'. PCR product was inserted into pDONR221 for propagation and sequencing and then transferred from pDONR211 to pDEST15. The pDEST15+SSI expression plasmid contains wild-type maize SSI cDNA which was expressed and purified from *E. coli* (see Figure 5).

A chimeric SS was created by fusing the GLASS and linker domains of GBSSI (amino acid residues 76-391 from Genbank accession no. X03035) with the SSI GLYTR domain (amino acid residues 376-642 of Genbank accession no. AF036891). This was accomplished with a two-step PCR reaction, in which portions of the insert were PCR amplified individually and then pieced together in a second PCR reaction. The GBSS region was amplified from the source plasmid pEXSwx2 using the of GLASS and linker domains identified in Commuri *et al.* [6] using the following primers: 1) 5' G GGG ACA AGT TTG TAC AAA AAA GCA GGC TTA CTG GTG CCA CGC GGT AGT ATG AAC GTC GTC TTC GTC GGC 3' and 2) 5' ACC TTC AGC AGT TGT GAC GTC CAC CGG GAG CCC GAC 3'. The SSI GLYTR domain [6] was amplified from pEXS1d using the following primers: 1) 5' GTC GGG CTC CCG GTG GAC GTC ACA ACT GCT GAA GGT 3' and 2) 5' GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TTA CAT GAC ATA GGG TCG 3'. PCR conditions for the first PCR were as follows: step 1 = 95°C for 10 min, step 2

= 95°C for 1 min, step 3 = 60°C for 1 min, step 4 = 72°C for 1 min, step 5 = repeat steps 2-4 35 times, and step 6 = 72°C for 5 min. The products from the first PCR were purified on a 1% agarose gel and excised on a UV light box. 50 µl of dH2O was added to each agarose piece and then homogenized with a pestle in a microfuge tube and incubated on ice for 1 h. Samples were centrifuged at 10,000g and the soluble portions of the solution containing the gel purified PCR fragments were used in the second PCR reaction. The second PCR reaction is called a "touch down PCR" and uses the following conditions: step 1 = 94°C for 2 min, step 2 = 94°C for 45 sec, step 3 = 72°C (-1°C every cycle) for 30 sec, step 4 = 72°C for 3 min, step 5 = repeat steps 2-4 12 times, step 6 = 94°C for 45 sec, step 7 = 60°C for 30 sec, step 8 = 72°C for 3 min, step 9 = repeat steps 6-8 12 times, and step 10 = 72°C for 2 min and contains the two gel purified pieces from the first PCR and the forward primer for GBSSI 1) 5' G GGG ACA AGT TTG TAC AAA AAA GCA GGC TTA CTG GTG CCA CGC GGT AGT ATG AAC GTC GTC TTC GTC GGC 3' and the reverse primer for SSI 2) 5' GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TTA CAT GAC ATA GGG TCG 3'. The final PCR product was inserted into pDONR221 for propagation and sequencing and then transferred from pDONR211 to pDEST15. The pDEST15+GBSSI-SSI expression plasmid contains maize GBSS-SSI cDNA which was expressed and purified from *E. coli* (see Figure 6).

Two point mutants were created by PCR mutagenesis that altered the amino acid sequence of the linker region in SSI. The first one, named SSI point mutation (PM) 1, introduced two point mutations in close proximity to one another near the N-terminal portion of the linker region [6]. One mutation generated a C to T transition that would result in a shift from an arginine to an alanine at position 1464 in the peptide. This mutation is predicted to rid the linker of a basic bulky side chain and replace it with a simple amino acid. The second mutation changed a C to a G that would result in a shift of a proline to a glycine which would cause the linker region to become less rigid as glycine is more flexible than a proline. These mutations were also created using two-step PCR that required amplification of the first fragment and use of that fragment as the 5' primer for the second reaction. PCR

conditions for the first PCR reaction were as follows: step 1 = 95°C for 10 min, step 2 = 95°C for 1 min, step 3 = 60°C for 1 min, step 4 = 72°C for 3 min, step 5 = repeat steps 2-4 35 times, and step 6 = 72°C for 5 min. The first fragment was PCR amplified from pEXS1d using the upstream oligonucleotide 5' G GGG ACA AGT TTG TAC AAA AAA GCA GGC TCC CTG GTG CCA CGC GGT AGT ATG AGC ATT GTC TTT GTA ACC 3' and the downstream oligonucleotide 3' CCT CCT GCG CCA TTC ACC GAA TAC CCA CTC CAG 5'. The product from the first amplification was gel purified using Qiagen's gel extraction kit and then that fragment was used as the upstream primer for the second PCR along with the downstream oligonucleotide 3' GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TTA CAT GAC ATA GGG TCG 5' using the following PCR conditions: step 1 = 95°C for 10 min, step 2 = 95°C for 1 min, step 3 = 60°C for 1 min, step 4 = 72°C for 3 min, step 5 = repeat steps 2-4 35 times, and step 6 = 72°C for 5 min. PCR product was inserted into pDONR221 for propagation and sequencing and then from pDONR211 to pDEST15. The pDEST15+SSI-P1445G,R1464A expression plasmid contains maize SSI cDNA with 2 point mutations which was expressed and purified from *E. coli* (see Figure 7).

The second linker region mutation, SSI-G1488P, was created using the methods described above for SSI-P1445G,R1464A. In this case, the mutation caused a G to C transversion at position 1488 that would result in a shift from a glycine to a proline, causing the linker region to become more rigid. The first fragment was created from the pEXS1d template using the upstream oligonucleotide 5' G GGG ACA AGT TTG TAC AAA AAA GCA GGC TCC CTG GTG CCA CGC GGT AGT ATG AGC ATT GTC TTT GTA ACC 3' and the downstream oligonucleotide 3' CAA AAA ATT AAC TGC CTC AGG CTT GTC AAG GGC ATG CCT 5'. This fragment was used as the upstream primer for the second PCR along with the downstream primer 5' GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TTA CAT GAC ATA GGG TCG 3'. PCR product was inserted into pDONR221 for propagation and sequencing and then from pDONR211 to pDEST15. The pDEST15+SSI-G1488P expression plasmid contains maize SSI cDNA with 1 point mutation which was expressed and purified from *E. coli* (see Figure 8).

Protein purification and sequencing

The five constructs containing sequence for SS enzymes mentioned above were expressed in the *E. coli* Tuner strain (Novagen) because they enable adjustable levels of protein expression throughout all of the cells in a culture. After expression and cell lysis via sonication, the soluble GST-tagged fusion proteins were purified using affinity chromatography with glutathione sepharose™ 4B beads. After purification, GST-tag was removed from the SS enzyme by thrombin cleavage. Thrombin and cleaved GST-tag was removed from the purified SS enzyme in order to perform the activity assay and affinity electrophoresis without interference. The purification scheme for each protein was run on SDS-PAGE and was Coomassie stained or labeled for immunodetection.

The pDEST15+SSI purification scheme is a good example of how the purification process worked (see Figure 9). The Coomassie stained gel (gel A) shows the entire scheme starting with the supernatant after the cells were sonicated (lane 2) which contains multiple proteins from the *E. coli* extract including the GST-fusion protein with SSI. The sonicate was incubated with glutathione sepharose™ 4B in which the GST-fusion protein should bind. Lane 3 contains a small sample of those beads with the bound protein which can be seen around 83 kD, the predicted size of SSI with GST-tag. Lane 4 is the flow through of sonicate off of the glutathione sepharose™ 4B column which contains all of the bacterial expressed proteins minus the GST-fusion protein because most of it should be bound to the beads. The column was washed three times with 15 ml PBS and lane 5 is a sample from the final wash and does not contain any detectable protein indicating that the washes were successful. The bound GST-fusion proteins were eluted off of the beads using reduced glutathione. The eluted proteins were concentrated and loaded in lane 6 (10 μg/ml). Only a single band of the GST-fusion protein, 83 kD, should be seen but another significant band is seen at approximately 60 kD. This band is later identified using N-terminal amino acid sequencing and immunodetection and is discussed later. Next, the GST-tag is cleaved off of the SSI enzyme using the incorporated thrombin cleavage site. The cleaved GST-tag is

removed from the SSI enzyme by the addition of more glutathione sepharose™ 4B. The GST-tag binds to the beads leaving the SSI enzyme and thrombin in solution. The GST-tag bound protein to the glutathione beads are found in lane 7 but the GST-tag band of approximately 25 kD can not be seen because it is too small and has run off the gel. A faint band can be seen in this lane and may be present if the beads were not washed well enough. The biotinylated thrombin is removed from the SSI enzyme by the addition of streptavidin agarose. The thrombin binds the beads leaving the pure SSI enzyme in solution. Lane 8 contains the streptavidin beads with the bound thrombin and no bands should be seen. However, a couple of faint bands are visible and may be present if the beads had not been washed well before loading onto the gel. The purified SSI enzyme was loaded in lane 9 and only a single band of 56 kD (predicted size of this SSI enzyme) should be seen but another band of approximately 60 kD is also seen. This 60 kD protein is the same "extra" protein seen in lane 6.

Immunodetections of the protein purification scheme using a similar gel line up as the Coomassie stained gel of the SSI purification can also be found in Figure 9. Gel B was detected by protein gel blot analysis using the probe α-SSI (1:1000, received from Dr. Guan and Dr. Keeling, BASF-Plant Sciences, Ames, IA) and SSI is identified across the gel at the correct predicted sizes mentioned above in the Coomassie gel explanation. Gel C was probed with α-GST (1:4000, BP Biosciences Pharmingen) and the GST-fusion protein at 83 kD can be detected in lanes 2-4, 6, and 9. GST detection in lane 4 indicates that not all of the fusion protein was bound to the glutathione beads. Lane 9 detection shows that not all of the GST-tagged SSI was removed from the purified fraction.

An interesting outcome of the protein purification process was the induction and co-purification of a second protein along with the protein of interest (see Figures 9-14). This protein's identity was unknown and was submitted for N-terminal amino acid sequencing from the co-purification with GBSSI (see extra band around 60 kD in Figure 10, gel A, lane 9) resulting in the amino acid sequence KDVKFGND. BLAST [19] analysis revealed a high correlation to GroEL protein, 57 kD, from *E. coli* which is a chaperonin that aids in protein

folding [20]. To make certain that this protein was co-purifying with the SS enzyme instead of having affinity to the glutathione beads, another purification scheme of "empty" Tuner cells was compared to pDEST15+GBSSI-SSI purification scheme (see Figure 14). No significantly enhanced band at 57 kD appeared in the "empty" tuner cells but did appear in the pDEST15+GBSSI-SSI. Immunodetection of the same SSI enzyme purification scheme mentioned above probed with α-GroEL (1:4000, created in our laboratory by Tracie Bierwagen) indicated that the protein that co-purified with the SSI enzyme was in fact GroEL (see Figure 9, gel D).

Similar expression patterns to the SSI enzyme can be seen for the remaining SS enzymes. pDEST15+GBSSI purification scheme is seen in Figure 10 and the predicted size of the GST fusion with GBSSI is 80 kD (lane 6) and the predicted size of the purified GBSSI enzyme is 55 kD. Gel B was detected by protein gel blot analysis using the probe α-GBSSI (1:4000, received from Dr. Guan and Dr. Keeling, BASF-Plant Sciences, Ames, IA) and GBSSI is identified across the gel at the correct predicted sizes. Gel C is probed with α-GST and the GST-fusion protein was detected in all of the lanes except lane 5 which is the final column wash. All of the GST-fusion protein must not have bound to the glutathione beads in lane 3 and in lanes 7-9 a residual amount of protein was present but was not even visible in the Coomassie stain.

pDEST15+GBSSI-SSI purification scheme can be seen in Figure 11. The predicted size of the GST fusion with GBSSI-SSI is 91 kD (lane 6) and with the GST-tag cleaved off of the enzyme it should be 64 kD (lane 9). Gel B was probed with α-GBSSI (1:4000) and the chimeric protein was identified in all lanes except lane 5 (final column wash). The band in lane 4 indicates that not all of the protein was bound to the beads and a residual amount of the protein is seen in lanes 7 and 8. Gel C was probed with α-SSI (1:1000) and the most prominent bands are seen in lane 3 (beads), 6 (GST fusion with GBSSI-SSI), and 8 (streptavidin beads). The band seen in lane 8 is somewhat of a mystery but is usually seen due to residual protein. However, a small amount is still visible in lane 9 which contains the purified protein. Gel D is probed with α-GST (1:4000) and bands are visible in lanes 2-4, 6,

and 9. All of the GST-fusion protein must not have bound to the glutathione beads in lane 3 and in lane 9 a residual amount of protein was present but was barely visible in the Coomassie stain.

pDEST15+SSI-P1445G,R1464A and pDEST15+SSI-G1488P purification schemes can be see in Figures 12 and 13 respectively. These two schemes are very similar because the proteins being investigated are very similar. The predicted size of the GST fusion with the SSI enzyme (PM1 and PM2) is 83 kD (lane 6) and with the GST cleaved off it should be 56 kD (lane 9). Gel B in both figures was probed with α-SSI (1:1000) and the most prominent bands are seen in lane 3 (beads), 6 (GST fusion with SSI (PM1 or PM2)), and 9 (purified protein without GST-tag). Gel C in both figures was probed with α-GST (1:4000) and bands are visible in lanes 2 (supernatant), 3 (beads), and 6 (GST fusion with SSI (PM1 or PM2)). Gel D in both figures was probed with α-GroEL (1:4000) and bands were detected in all lanes except lane 5 which is the final wash of the column.

Zymogram analysis

Zymogram analysis was a method used to visualize the activity of SS enzymes on rabbit liver glycogen. Purified SS enzymes were denatured in SDS loading buffer by boiling for 10 min. Proteins were loaded onto an 8% acrylamide SDS-PAGE gel containing 0.1% rabbit liver glycogen. Because the SS enzymes are denatured, the enzymes should not interact with the glycogen in the gel which could cause a reduction in relative mobility. After separation, the gel was washed to remove any residual SDS that could prevent the enzymes from renaturing. Then the gel was incubated in a reaction buffer which contained all of the ingredients necessary for an SS enzyme to act on a substrate. An active SS enzyme with an affinity for glycogen will elongate the glycogen's glucan chains and which is detectable by iodine staining.

Analysis of the purified SS enzymes is seen in Figure 15. Dark bands indicate an elongation of the glycogen's glucan chains and can be seen in lanes 2 (SSI), 4 (SSI-

P1445G,R1464A), and 5 (SSI-G1488P). No visible activity was detected in the wild-type GBSSI (lane 1) or the chimeric protein GBSSI-SSI (lane 3). However, this type of activity assay has not shown success with GBSSI activity most likely due to the lack of a semi-crystalline substrate to act upon.

Quantification of activity

Quantification of SS enzyme activity can be investigated using the incorporation of radiolabeled ADPG into a glucan substrate [13]. In this experiment, two different glucan substrates were used to investigate the purified enzymes' activity including amylopectin and rabbit liver glycogen. The hope was that the amylopectin would provide the semi-crystalline structure needed for GBSSI to be active. A liquid scintillation counter was used to detect the amount of radiolabeled ADPG into either amylopectin or rabbit liver glycogen.

Purified wild type SSI is the only enzyme that exhibited a significant amount of activity (see Figure 16). A slight increase in activity of GBSSI-SSI can be seen with the amylopectin substrate. No significant activity of the wild type GBSSI is detected with either substrate.

Affinity electrophoresis

Affinity electrophoresis is a method used to determine an enzymes affinity for a particular substrate [11]. In this study, the purified SS enzymes' affinity for rabbit liver glycogen was investigated. Purified proteins were loaded onto native 6% acrylamide PAGE gels with varying concentrations of rabbit liver glycogen (0%, 0.5%, 1.0%, 1.5%, 2.0%, and 2.5%). The gels were Coomassie stained and migration distances of the protein bands were measured.

Only those proteins whose relative mobility was slowed due to interaction with the rabbit liver glycogen substrate, including wild-type SSI and SSI-G1488P, remained on the gels and could be measured (see Figures 17 and 18 (lanes 3 and 6)). Notice the increase of affinity to the gel, shorter relative mobility, with the increase of concentration of rabbit liver

glycogen with these two SS enzymes. Wild-type GBSS, GBSS-SSI, and SSI-P1445G,R1464A all ran off of the gels and their affinity, if any, to the rabbit liver glycogen was not enough to be measured using this method.

Discussion

Important goals in the understanding of the functionality of maize GBSSI and SSI include investigation of the function GLASS and GLYTR domains which make up the catalytic domain in all SS's [6]. A total of five SS proteins were cloned (see Figures 4-8) and then expressed and purified in *E. coli* and the apparent molecular mass of the recombinant protein matched that predicted by the primary sequence (see Figures 9-13). GBSSI and SSI enzymes were controls used for the comparison to the chimeric and mutagenized enzymes. The chimeric protein consisting of a fusion of the GLASS domain of GBSS with the GLYTR domain of SSI, GBSSI-SSI, was used to investigate the functionality of the GLASS and GLYTR domains of SS's. Two mutagenized SS enzymes with point mutations in the linker region between GLASS and GLYTR were used to investigate the consequences of recombination in this region.

Each recombinant protein was examined for SS activity in SDS-PAGE activity gels (starch synthase zymograms). The zymogram analysis shows that wild-type SSI, and the two mutant proteins SSI-P1445G,R1464A and SSI-G1488P, possessed the enzymatic activity to elongate the chains of the rabbit liver glycogen substrate. Even though this is not a quantitative test, it appears that the wild-type SSI had more activity than the two mutant proteins. It also appears that SSI-G1488P has a little more activity than SSI-P1445G,R1464A. In contrast, the absence of SS activity bands in the zymogram analyses of the wild-type GBSSI and chimeric GBSSI-SSI proteins does not necessarily indicate a loss of activity in those enzymes but the inability of these enzymes to act on a glycogen substrate. GBSSI has been shown to require a crystalline substrate for activity, such as might be found

within a starch granule, and that it is not possible or very difficult to duplicate these conditions *in vitro* [13].

One explanation of the finding that the GBSSI-SSI chimeric protein did not show any activity even though it contained the two necessary domains required of a starch synthase to operate is that GLASS and GLYTR cannot act independently of one another to elongate glucan chains. The function of the GLYTR region is suggested to be involved in interaction with the glucan substrate during the process of glucan chain elongation. If the GLYTR region was able to act independently of the GLASS region then the enzyme should have been able to elongate the glycogen substrate. This suggests that the use of the GBSS GLASS domain in the chimeric enzyme makes it behave in a manner similar to GBSS with the respect to glucan binding requirement (GBSS needs a semi-crystalline structure and a supply of MOS [13]), and that the glycogen substrate provided in the gel is not an appropriate substrate for binding. This supports the hypothesis the interaction of the GLASS and GLYTR domains have an effect on an enzyme's degree of functionality to act on a substrate. Further examination of this explanation could be investigated by making another chimeric protein fusing the SSI GLASS domain and the GBSS GLYTR domain and testing its activity in an SDS-PAGE activity gel containing rabbit liver glycogen.

Another possible explanation why the chimeric protein did not show any activity may be that the enzyme is not folded properly. Even though the GLASS and GLYTR domains of all SS enzymes are similar, specific orientation of those two domains may be required in order for an enzyme to be active. Determining the exact structure of SS enzymes and comparing them to the structure of the chimeric SSI enzyme by X-ray crystallography would be a way to determine structural differences, if any.

Quantification of the activity of the SS enzymes GBSSI, SSI, and GBSSI-SSI was investigated using the incorporation of radiolabeled ADPG into a glucan substrate [13]. Two different glucan substrates were used to investigate the purified enzymes' activity including amylopectin and rabbit liver glycogen with the hope was that the amylopectin would provide the semi-crystalline structure needed for GBSSI to be active. However, the experiment only

showed significant activity of the SSI enzyme to act on the amylopectin and the glycogen substrates (see Figure 16). Perhaps more than just the presence of amylopectin is required for the GBSSI enzyme to be active. The same explanations why the chimeric showed little to no activity in the zymogram analysis apply to why the chimeric showed little to no activity in this experiment. This assay may also be improved with the use of the Dowex bead method [13].

The binding efficiency of each SS to a branched glucan substrate was tested by affinity electrophoresis. The binding efficiency based on affinity electrophoresis showed that wild-type SSI and SSI-G1488P had an affinity to the rabbit liver glycogen substrate in the gel (see Figures 17 and 18). Wild-type GBSS, GBSS-SSI, and SSI-G1488P all ran off of the gel which shows little to no interaction with the substrate. However, because the proteins were run under native conditions, the separation is based on charge or conformation opposed to molecular weight. Because the pH of the running buffer was 8.8 it would not seem that the proteins did not enter the gel due to the loss of charge (see Table 2) which happens when the pH is equal to the pI of the protein. Perhaps the proteins that were not visible on the gel were not folded properly causing it to run different from the SSI and SSI-G1488P.

It is also interesting to note that SSI and SSI-G1488P have similar K_d's according to the graph in figure 17 because the slopes of the lines are similar. However, SSI-G1488P has approximately 15% lower affinity compared to SSI. By only changing one amino acid, the affinity for the substrate is changed indicating that the linker region cannot be altered without having an effect on the enzyme's functionality.

This region of the enzyme appears to play an essential role in defining the enzymatic properties of SSs. A repeat of the affinity electrophoresis of these proteins needs to be performed in order to confirm these observations. Different substrates could also be used, ADPG or amylopectin, to investigate any further differences in enzymatic properties. Additional investigation of the enzymatic properties of these proteins could also be tested by using the radioactive uptake assay with a variety of substrates.

The question of the importance of the linker region to SS function was tested by altering specific residues within the linker, such that this region would become either more flexible (PM1) or less flexible (PM2). Zymogram analysis of both point mutant proteins showed they exhibited SS activity similar to that of the wild-type SSI enzyme on the gel. However, the affinity electrophoresis indicated otherwise as mentioned above. This implies that the introduced changes to the linker region did result in alteration of either enzymatic activity or glucan binding properties. A certain level of rigidity of an SS enzyme may be required for an enzyme to bind to a substrate. This hypothesis that rearrangement in the linker region of SSI does not cause a loss of activity of the enzyme held true even though glucan binding was altered.

A future application of this research would involve creating starch with altered chain length distribution of industrial interest using chimeric SS's. Also if the predicted function of the GLASS domain is correct, then construction of fusion proteins with the GLASS region of GBSS may result in the entrapment of foreign proteins in the granule which would result in starch with added value. These are just a couple of directions in which this research could go to create novel starch and further the understanding of SS enzymes.

Another direction of SS research may involve testing the interaction between SS enzymes and the GroEL chaperonin. In all the SS expressions in this research, GroEL was detected in the final purified fraction. If the GroEL protein was co-purifying based on its affinity to the glutathione sepharose™ 4B then it would have been removed from the final purified protein when the cleaved GST-tag was removed. However, the GroEL demonstrated an association to the SS enzyme by its retention in the final purified fraction. To be sure that this result was true, a repeat of expression and purification of the chimeric protein and "empty" *E. coli* cells was performed (see Figure 14). The result of this showed that Hsp60 did not bind to the glutathione sepharose™ 4B in the "empty" *E. coli* cells but did bind with the chimeric protein indicating an interaction between the SS enzymes being studied and GroEL. A similar interaction with GroEL has been shown in the purification of branching

enzyme IIa (BEIIa), *su1* which codes for an isoamylase type debranching enzyme (DBE), and *zpu1* which codes for a pullulanase type DBE [21-23].

GroEL is a member of the heat shock protein 60 (Hsp60) class of chaperones and has been shown to aid in protein folding when a protein has a problem with aggregation before it can finish being folded properly [20]. Chaperones are ubiquitous and are required for folding of some other proteins including ribulose bisphosphate carboxylase (Rubisco) [24]. Perhaps SS enzymes need assistance with protein folding due to problematic aggregation before it can be suitably folded when being expressed in *E. coli*. There is evidence *in vivo* that proteins entering an amyloplast may need assistance with folding. Yu *et al.* [25] found an 81 kD protein associated with maize amyloplasts stroma which is a member of the Hsp70 family of stress-related proteins. Hsp70's function as molecular chaperones and their presence is found not only in the amyloplasts but also in the chloroplasts and chromoplasts. This discovery may help explain how polypeptides are translocated across the amyloplast envelope and how proteins are folded within the stroma [25]. This would also help explain the interaction of *E. coli* GroEL with SS's. If SS's need assistance with folding *in vivo*, then it is not surprising that they would need assistance *in vitro*.

This study illustrated that the interaction between the GLASS and GLYTR domains effects the degree of functionality of the enzyme. It also demonstrated that the linker region between the GLASS and GLYTR domain can not be modified and still retain enzymatic properties indicating the important of the linker region of SS enzymes. An interesting result from this research was the interaction of GBSS and SSI enzymes with the GroEL chaperone which may indicate that SS enzymes are folded *in vivo* with a chaperone protein of some type. However, further research is required to explain the precise function of each SS domain in order to understand how exactly how each individual SS enzyme functions.

References

1. Institute, I.S., *Starch Application Table*. 1999: Science Park Aarhus, Denmark.

2. James, M.G., K. Denyer, and A.M. Myers, *Starch synthesis in the cereal endosperm.* Curr Opin Plant Biol, 2003. **6**(3): p. 215-22.

3. Nelson, O.E. and H.W. Rines, *The enzymatic deficiency in the waxy mutant of maize.* Biochem Biophys Res Commun, 1962. **9**: p. 297-300.

4. Shure, M., S. Wessler, and N. Fedoroff, *Molecular identification and isolation of the Waxy locus in maize.* Cell, 1983. **35**(1): p. 225-33.

5. Tetlow, I.J., M.K. Morell, and M.J. Emes, *Recent developments in understanding the regulation of starch metabolism in higher plants.* J Exp Bot, 2004. **55**(406): p. 2131-45.

6. Commuri, P.K., Peter L.; Ramirez, Nona; McKean, Angela; Gao, Zhong; Guan, Hanping, *Glucan Chain Length Domains*, in *United States Patent Application.* 2004.

7. MacGregor, E.A., *Possible structure and active site residues of starch, glycogen, and sucrose synthases.* J Protein Chem, 2002. **21**(4): p. 297-306.

8. Kelley, L.A., R.M. MacCallum, and M.J. Sternberg, *Enhanced genome annotation using structural profiles in the program 3D-PSSM.* J Mol Biol, 2000. **299**(2): p. 499-520.

9. Cao, H., et al., *Identification of the soluble starch synthase activities of maize endosperm.* Plant Physiol, 1999. **120**(1): p. 205-16.

10. Mu-Forster, C., et al., *Physical association of starch biosynthetic enzymes with starch granules of maize endosperm. Granule-associated forms of starch synthase I and starch branching enzyme II.* Plant Physiol, 1996. **111**(3): p. 821-9.

11. Commuri, P.D. and P.L. Keeling, *Chain-length specificities of maize starch synthase I enzyme: studies of glucan affinity and catalytic properties.* Plant J, 2001. **25**(5): p. 475-86.

12. Klosgen, R.B., et al., *Molecular analysis of the waxy locus of Zea mays.* Mol Gen Genet, 1986. **203**: p. 237-244.

13. Denyer, K., et al., *Interaction with amylopectin influences the ability of granule-bound starch synthase I to elongate malto-oligosaccharides.* Biochem J, 1999. **342 Pt 3**: p. 647-53.

14. Edwards, A., et al., *Specificity of starch synthase isoforms from potato.* Eur J Biochem, 1999. **266**(3): p. 724-36.

15. *Gateway Cloning Technology: Instruction Manual*, in *http://www.invitrogen.com/content.cfm?pageid=4072*. 2005.

16. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.

17. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.

18. Imparl-Radosevich, J.M., et al., *Purification and characterization of maize starch synthase I and its truncated forms.* Arch Biochem Biophys, 1998. **353**(1): p. 64-72.

19. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, 1997. **25**(17): p. 3389-402.

20. Walter, S., *Structure and function of the GroE chaperone.* Cell Mol Life Sci, 2002. **59**(10): p. 1589-97.

21. Kim, S. (unpublished data). Iowa State University, Ames, IA 50011.

22. Bierwagen, T.a.R., A. (unpublished data). Iowa State University, Ames, IA 50011.

23. Wu, C. (unpublished data). Iowa State University, Ames, IA 50011.

24. Goloubinoff, P., et al., *Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfoleded state depends on two chaperonin proteins and Mg-ATP.* Nature, 1989. **342**(6252): p. 884-9.

25. Yu, Y., et al., *Polypeptides of the maize amyloplast stroma. Stromal localization of starch-biosynthetic enzymes and identification of an 81-kilodalton amyloplast stromal heat-shock cognate.* Plant Physiol, 1998. **116**(4): p. 1451-60.

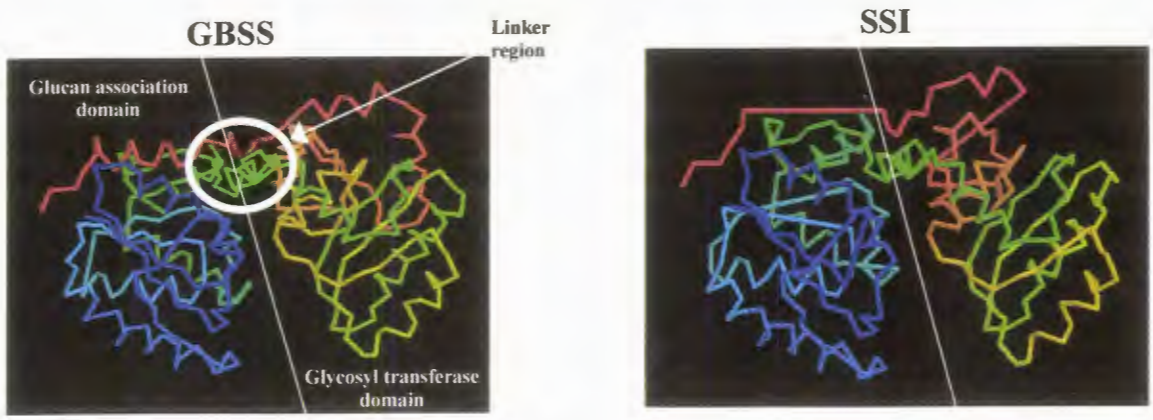


Figure 1: 3D-PSSM predicted structural models of GBSS and SSI [8].

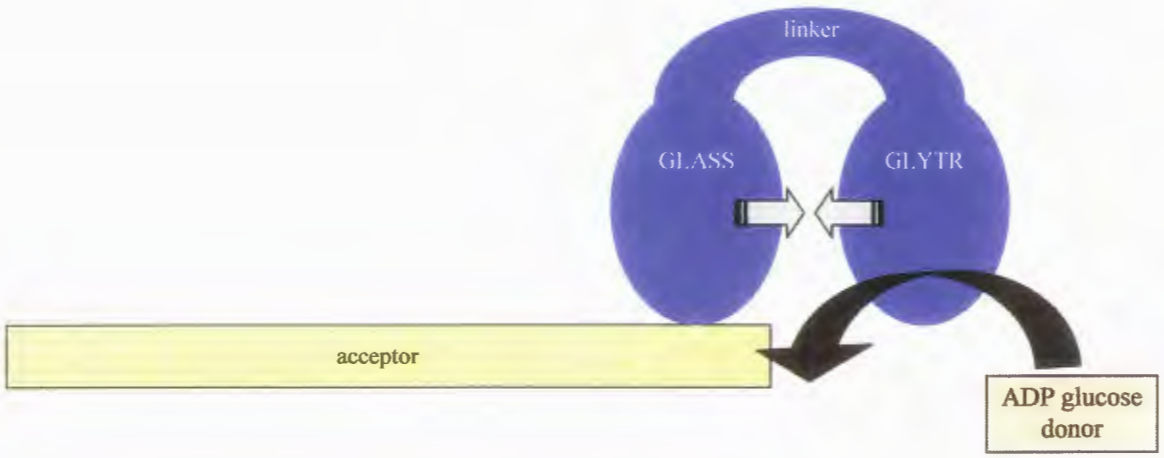


Figure 2: Schematic representation of the GLASS/GLYTR hypothesis.

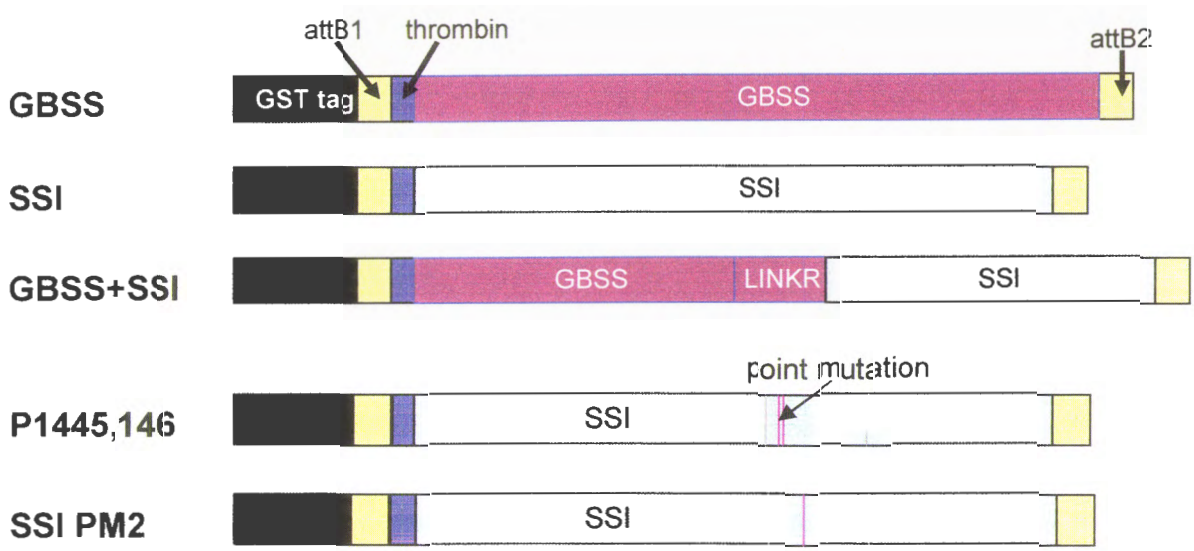


Figure 3: Visual representation of the five plasmids' construction.

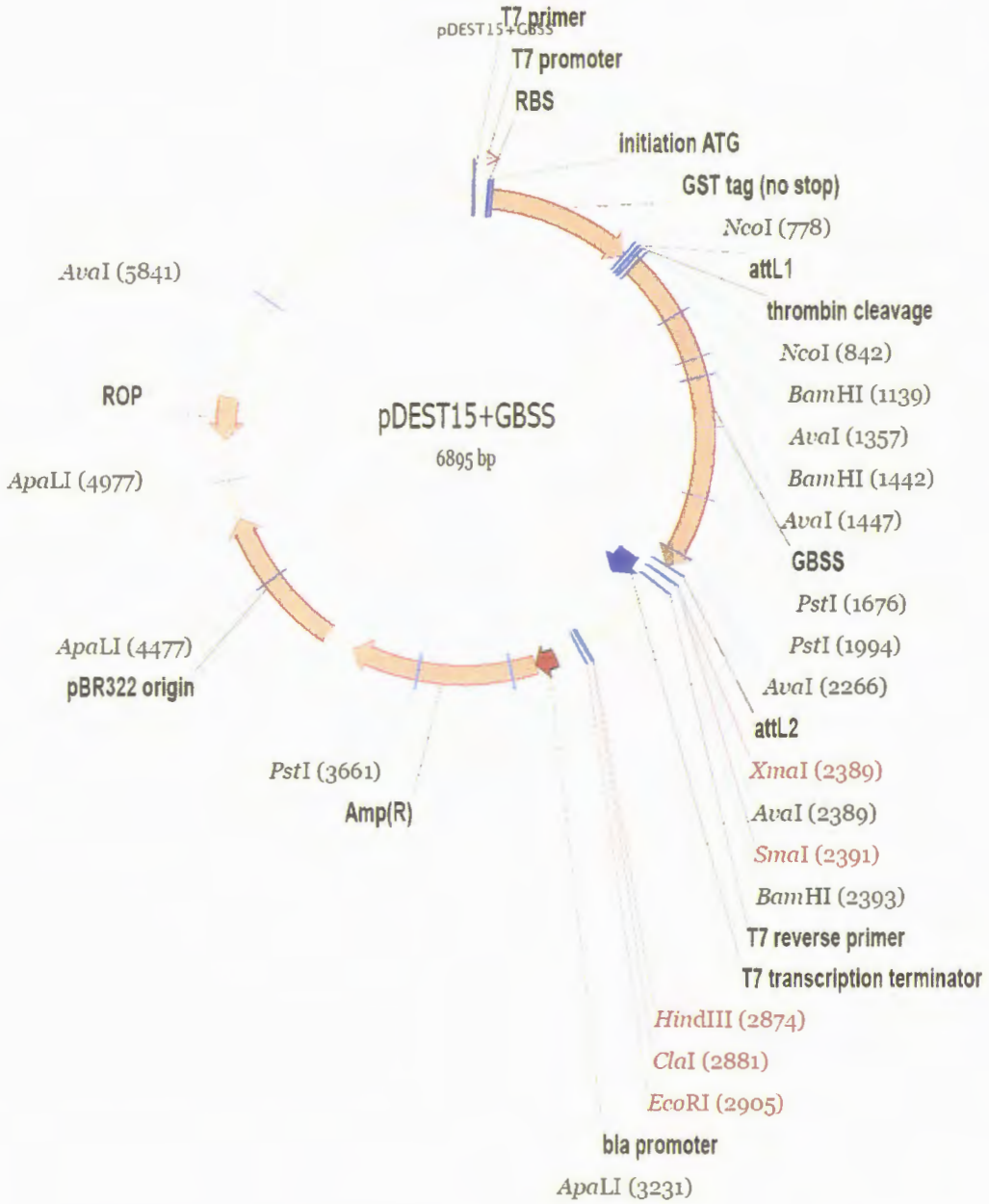


Figure 4: Plasmid map of the GBSSI cDNA insertion into the *E. coli* expression vector pDEST15 (Invitrogen). The GBSSI cDNA is comprised of an open reading frame from codon 840 to the native stop codon at 2426. Upstream of codon 840, in the pDEST15 expression vector, are the T7 promoter, a ribosome binding site (RBS), an ATG initiation codon, a 671bp of a GST-tag affinity purification sequence, a bacteriophage λ recombination site (attL), and a thrombin cleavage site.

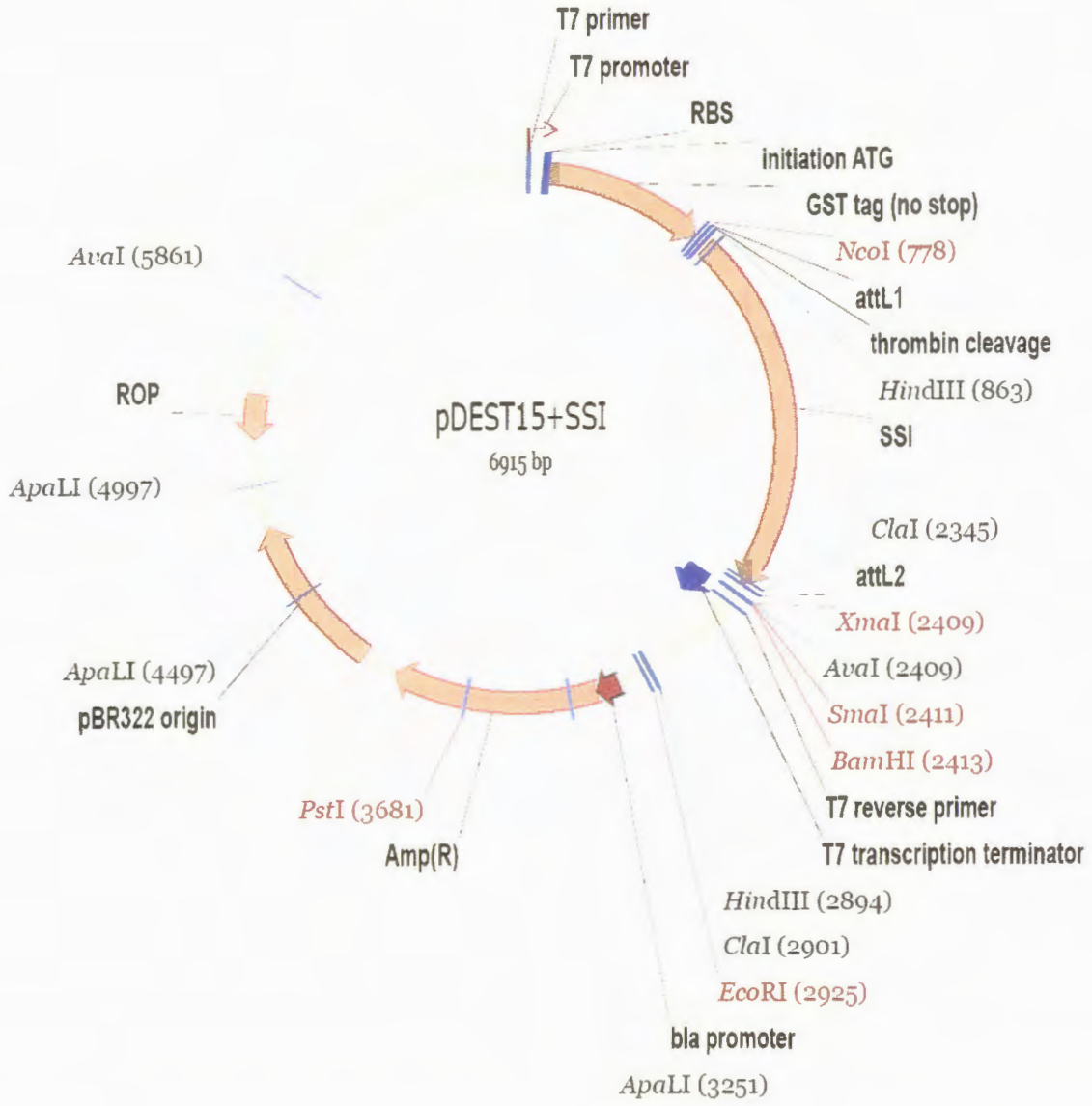


Figure 5: Plasmid map of the SSI cDNA insertion into the *E. coli* expression vector pDEST15 (Invitrogen). The SSI cDNA is comprised of an open reading frame from codon 840 to the native stop codon at 2366. Upstream of codon 840, in the pDEST15 expression vector, are the T7 promoter, a ribosome binding site (RBS), an ATG initiation codon, a 671bp of a GST-tag affinity purification sequence, a bacteriophage λ recombination site (attL), and a thrombin cleavage site.

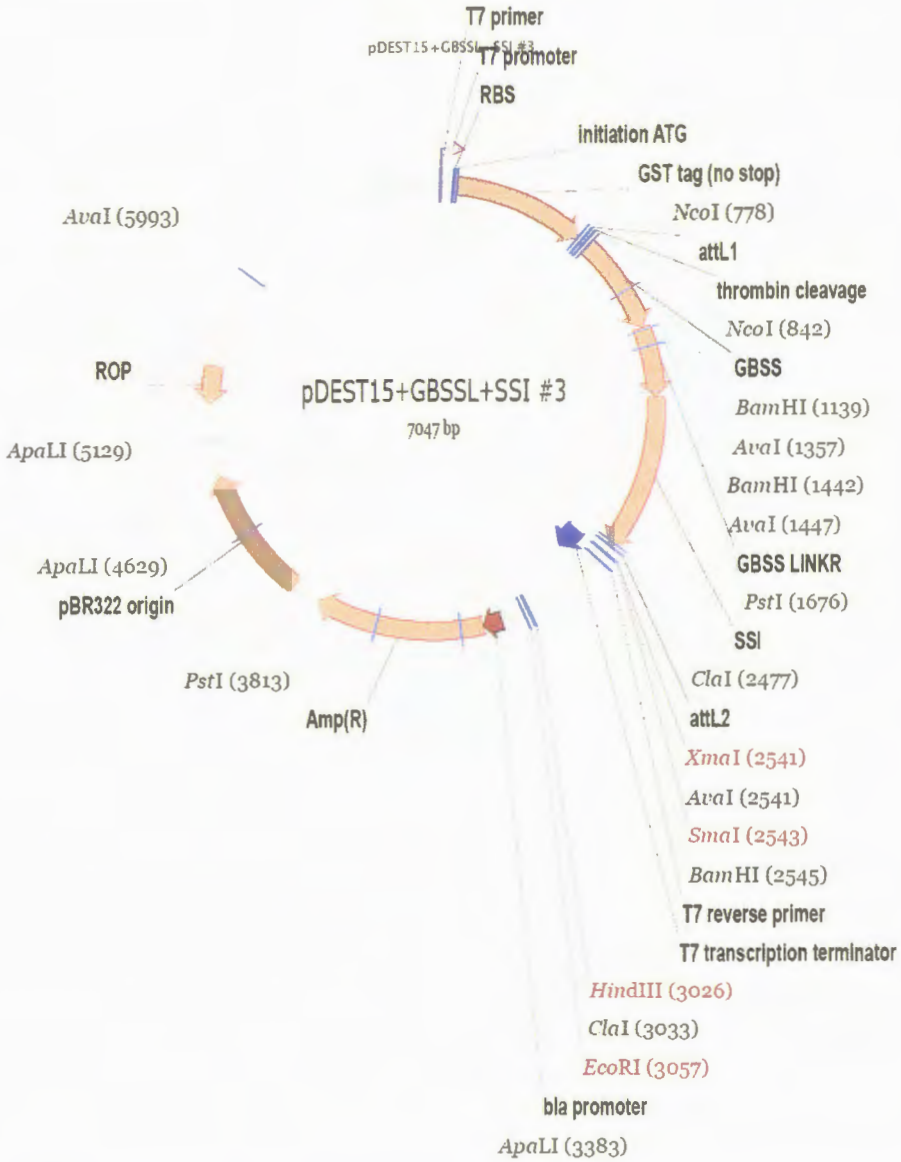


Figure 6: Plasmid map of the GBSSI-SSI cDNA insertion into the *E. coli* expression vector pDEST15 (Invitrogen). The GBSSI GLASS and linker cDNA is comprised of an open reading frame from codon 840 to codon 1781 and the SSI GLYTR cDNA is an open reading frame from codon 1782 to the native stop codon at 2579. Upstream of codon 840, in the pDEST15 expression vector, are the T7 promoter, a ribosome binding site (RBS), an ATG initiation codon, a 671bp of a GST-tag affinity purification sequence, a bacteriophage λ recombination site (attL), and a thrombin cleavage site.

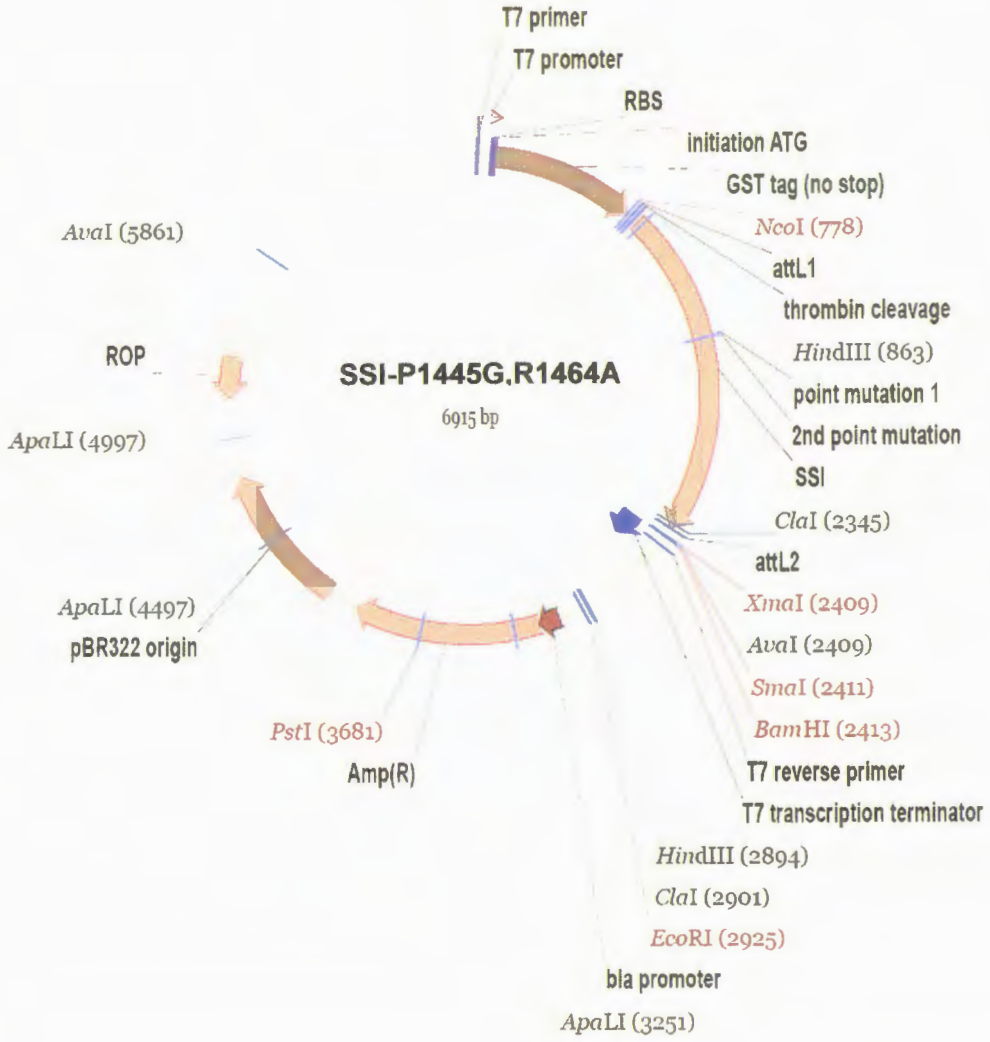


Figure 7: Plasmid map of the SSI-P1445G,R1464A cDNA insertion into the *E. coli* expression vector pDEST15 (Invitrogen). The SSI-P1445G,R1464A cDNA is comprised of an open reading frame from codon 840 to the native stop codon at 2366. Upstream of codon 840, in the pDEST15 expression vector, are the T7 promoter, a ribosome binding site (RBS), an ATG initiation codon, a 671bp of a GST-tag affinity purification sequence, a bacteriophage λ recombination site (attL), and a thrombin cleavage site. The first point mutation changing a proline to a glycine is located at codon 1455 and the second mutation changing a arginine to alanine is located at codon 1464.

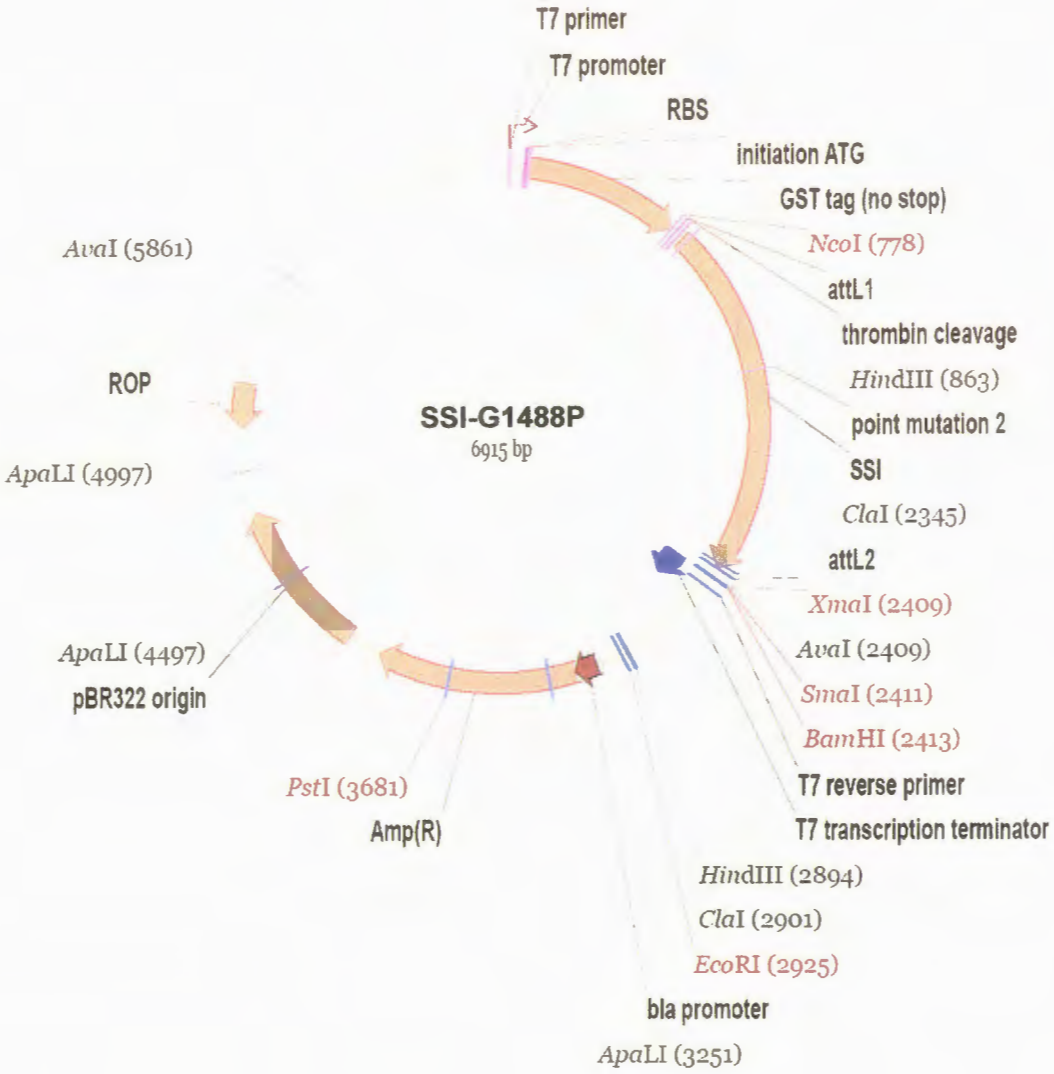


Figure 8: Plasmid map of the SSI-G1488P cDNA insertion into the *E. coli* expression vector pDEST15 (Invitrogen). The SSI-G1488P cDNA is comprised of an open reading frame from codon 840 to the native stop codon at 2366. Upstream of codon 840, in the pDEST15 expression vector, are the T7 promoter, a ribosome binding site (RBS), an ATG initiation codon, a 671bp of a GST-tag affinity purification sequence, a bacteriophage λ recombination site (attL), and a thrombin cleavage site. The point mutation is located at codon 1488.

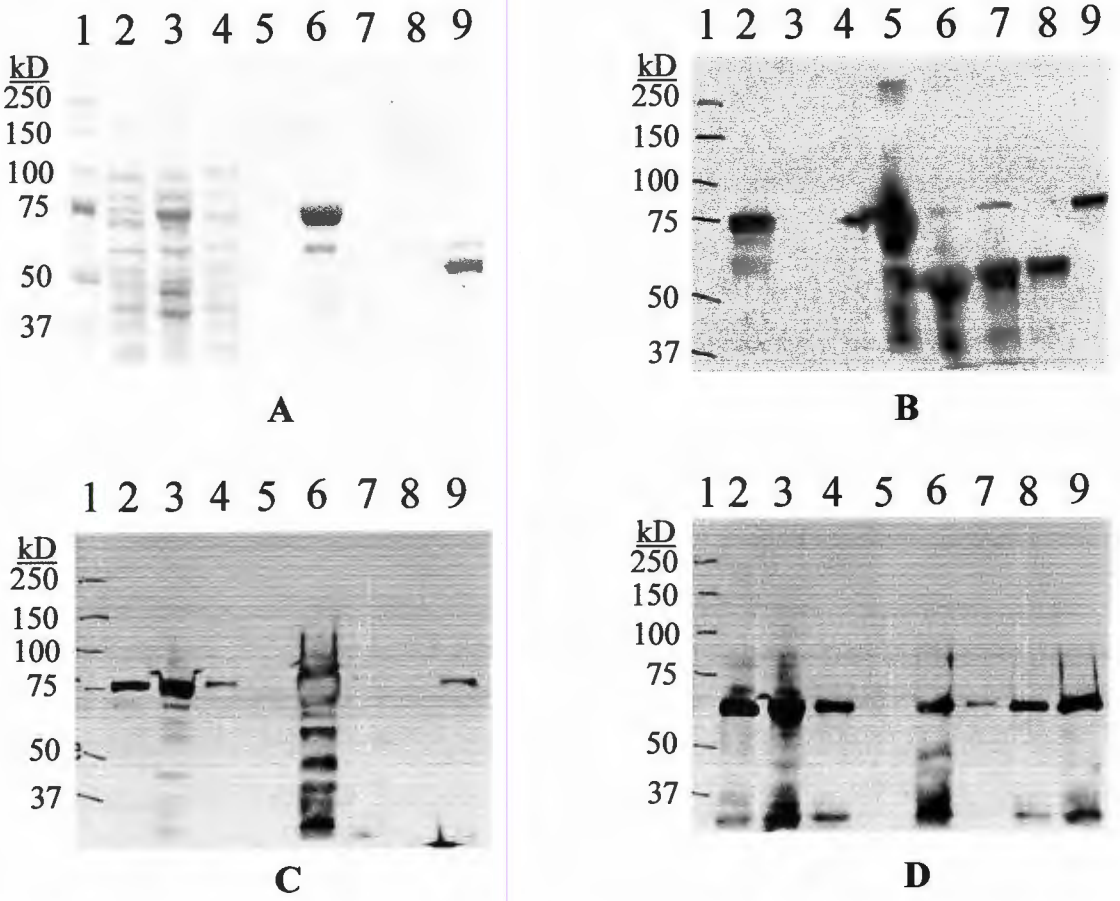


Figure 9: Purification scheme of SSI by SDS-PAGE. Gel (A) is Coomassie stained, (B) is an immunodetection with α-SSI, (C) is an immunodetection with α-GST, and (D) is an immunodetection with α-GroEL. (A), (C), and (D) have the following lane assignments: 1-molecular weight marker, 2-supernatant after sonication, 3-glutathione sepharose™ 4B with bound protein, 4-flow-through off column, 5-final wash, 6-eluted GST fusion protein off column, 7-glutathione sepharose™ 4B bound with thrombin cleaved GST-tag, 8-streptavidin agarose with bound biotinylated thrombin, 9-pure SSI protein without GST tag. (B) has the following land assignments: 1-molecular weight marker, 2- glutathione sepharose™ 4B with bound protein, 3-final wash, 4-first column elution of GST fusion protein, 5-third column elution of GST fusion protein, 6-pure SSI without GST tag, 7- glutathione sepharose™ 4B bound with thrombin cleaved GST-tag, 8-streptavidin agarose with bound biotinylated thrombin, 9-supernatant after sonication. The predicted size of SSI with the GST-tag was 83 kD, and without the GST-tag it was predicted to be 56 kD.

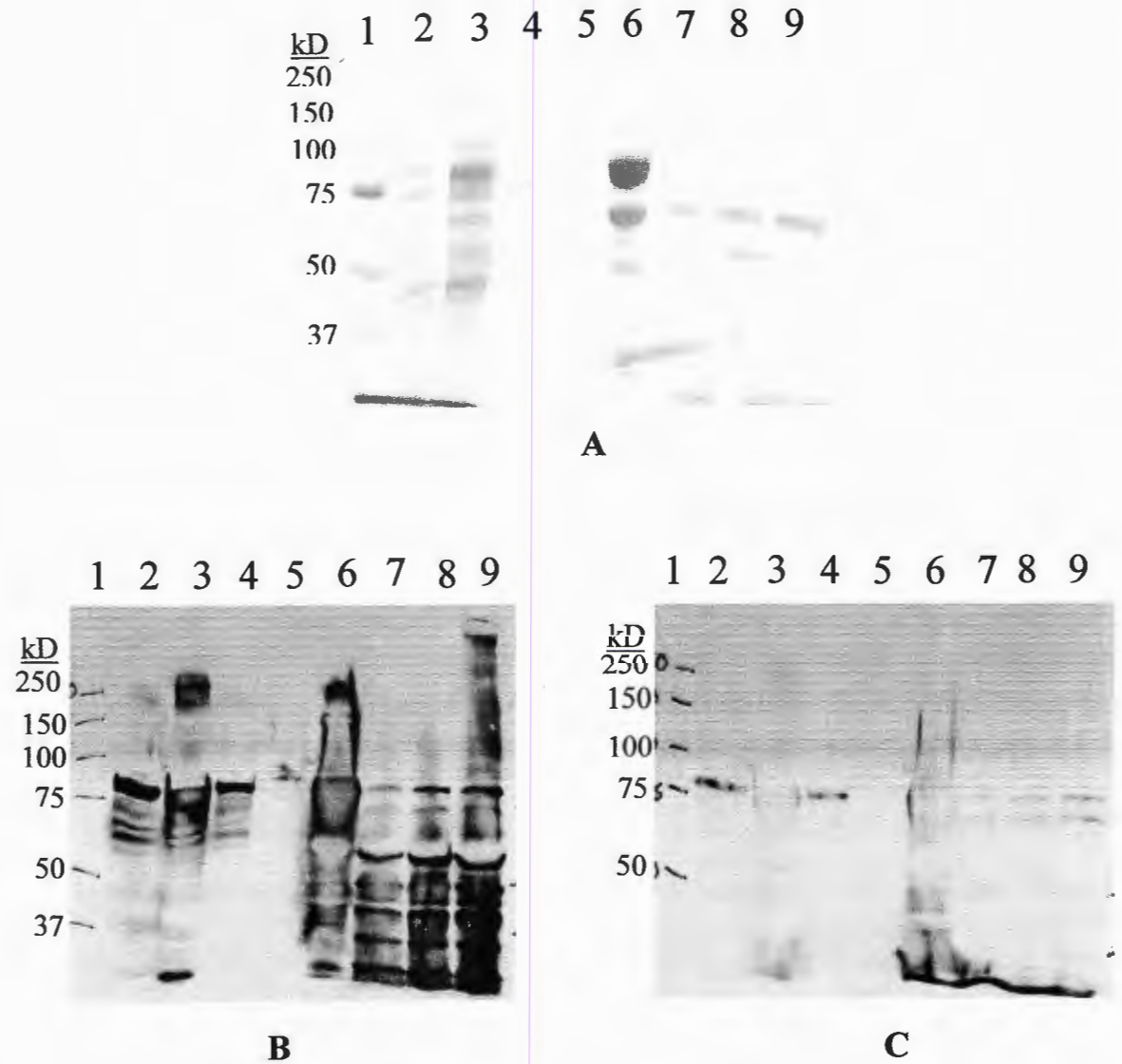


Figure 10: SDS-PAGE of GBSS purification scheme and verification by immunodetection. (A) is a Coomassie stained gel, (B) is an immunodetection with α-GBSS, and (C) is another immunodetection with α-GST. Lanes represent the following: 1-molecular weight markers, 2-supernatant after sonication, 3-glutathione sepharose™ 4B with bound protein, 4-flow-through off column, 5-final wash, 6-eluted GST fusion protein off column, 7-glutathione sepharose™ 4B bound with thrombin cleaved GST-tag, 8-streptavidin agarose with bound biotinylated thrombin, 9-pure GBSS protein without GST tag. GBSS with the GST-tag had a predicted size of 80 kD, and without the GST-tag it had a predicted size of 55 kD.

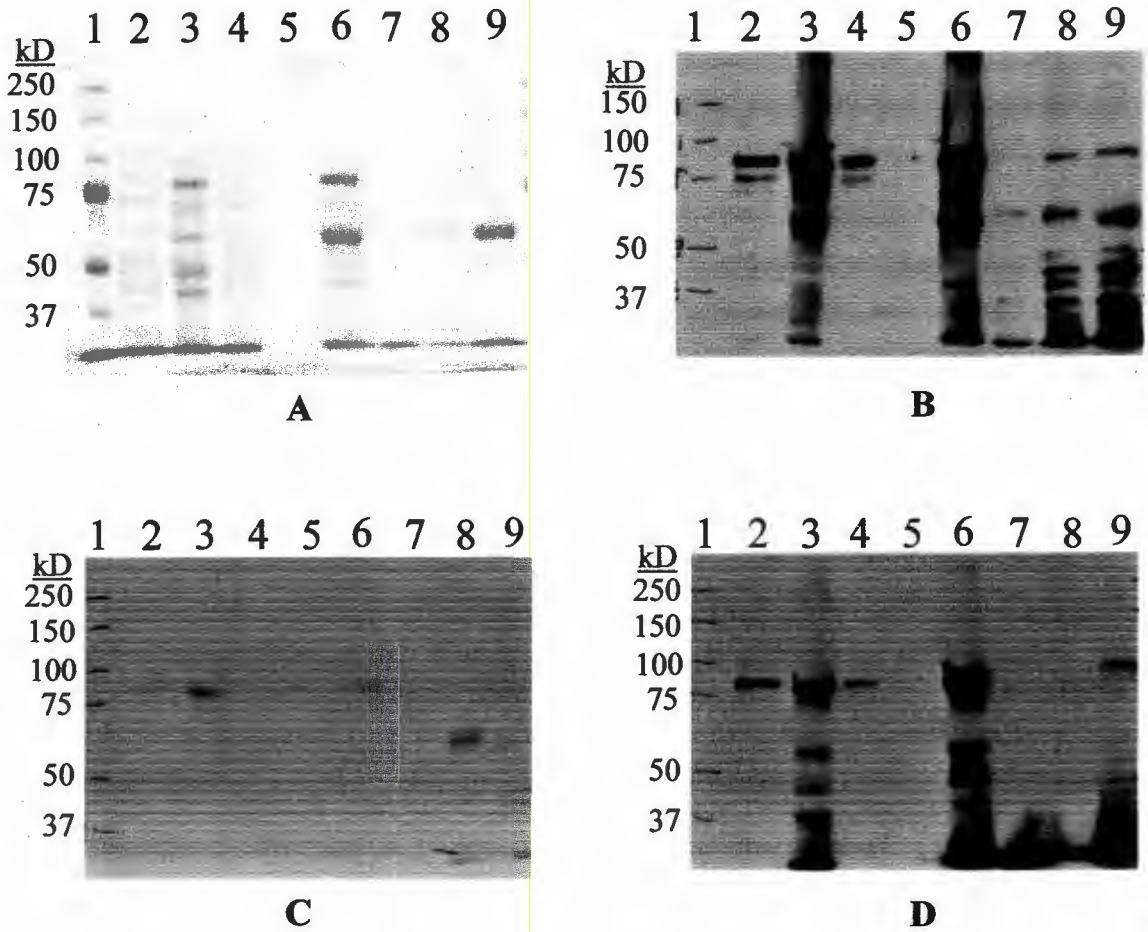


Figure 11: Purification scheme of chimeric protein GBSS-SSI. SDS-PAGE analysis of the chimeric protein includes (A) Coomassie stain, (B) immunodetection with α-GBSS, (C) immunodetection with α-SSI, and (D) immunodetection with α-GST. The following are the lane assignments for all four gels: 1-molecular weight markers, 2-supernatant after sonication, 3-glutathione sepharose™ 4B with bound protein, 4-flow-through off column, 5-final wash, 6-eluted GST fusion protein off column, 7-glutathione sepharose™ 4B bound with thrombin cleaved GST-tag, 8-streptavidin agarose with bound biotinylated thrombin, 9-pure GBSS-SSI protein without GST tag. The predicted size of GBSS-SSI with the GST-tag was 91 kD, and without the GST-tag it was predicated to be 64 kD.

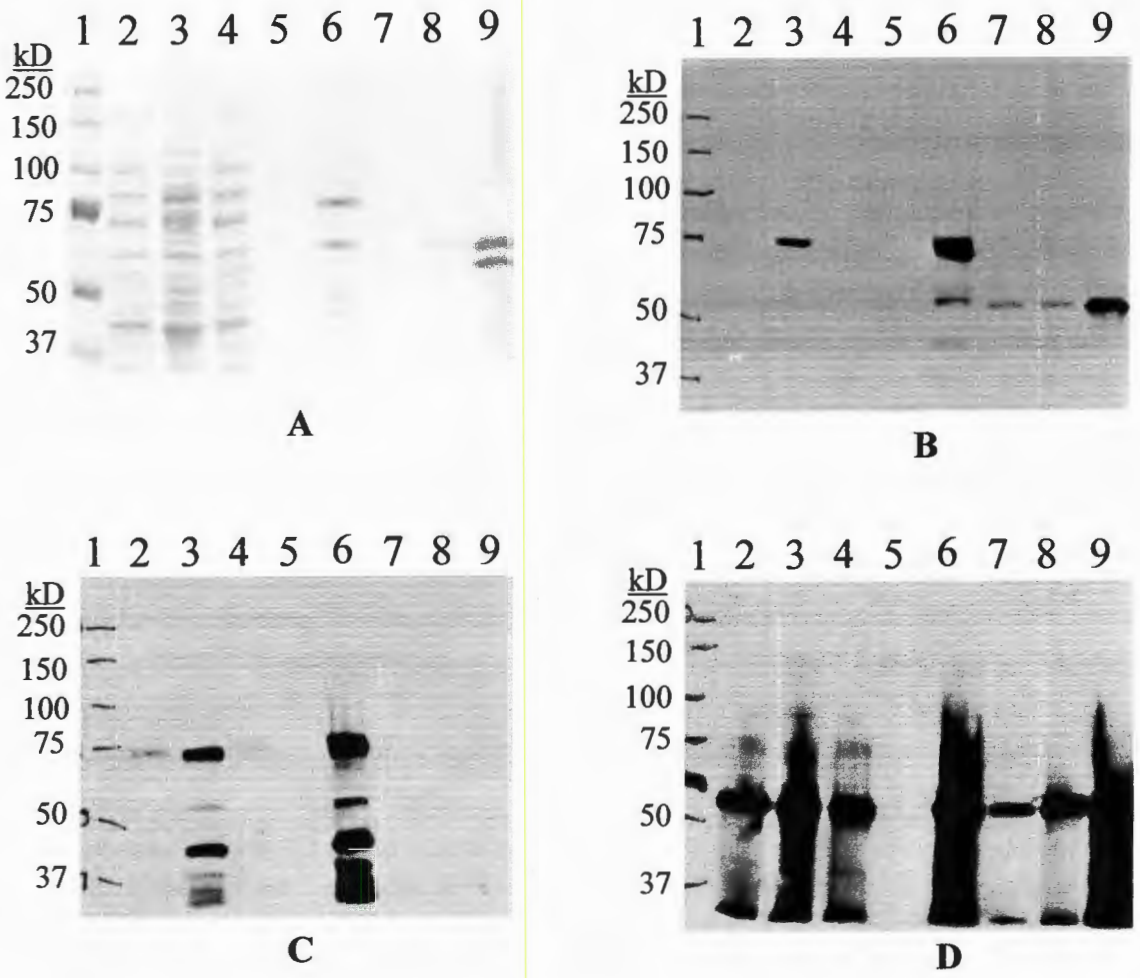


Figure 12: Purification scheme of point mutant SSI-P1445G,R1464A (2 point mutations). SDS-PAGE analysis by (A) Coomassie staining, (B) immunodetection with α-SSI, (C) immunodetection with α-GST, and (D) immunodetection with α-GroEL. The following are the lane assignments for all four gels: 1-molecular weight markers, 2-supernatant after sonication, 3-glutathione sepharose™ 4B with bound protein, 4-flow-through off column, 5-final wash, 6-eluted GST fusion protein off column, 7-glutathione sepharose™ 4B bound with thrombin cleaved GST-tag, 8-streptavidin agarose with bound biotinylated thrombin, 9-pure SSI-P1445G,R1464A protein without GST tag. The predicted size of SSI-P1445G,R1464A with the GST-tag was 83 kD, and without the GST-tag the predicted size was 56 kD.

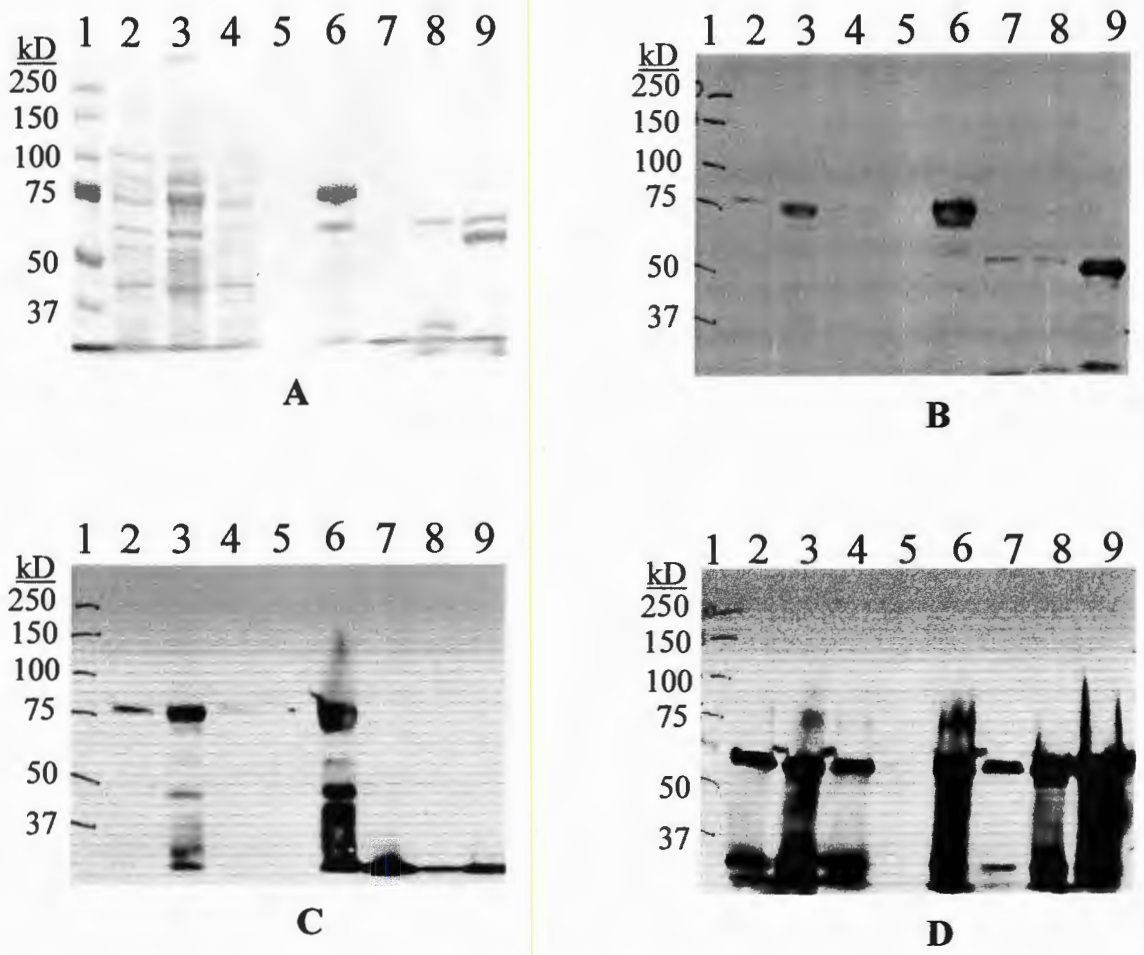


Figure 13: Purification scheme for point mutant SSI-G1488P (single point mutation). SDS-PAGE analysis by (A) Coomassie staining, (B) immunodetection with α-SSI, (C) immunodetection with α-GST, and (D) immunodetection with α-GroEL. The following are the lane assignments for all four gels: 1-molecular weight markers, 2-supernatant after sonication, 3-glutathione sepharose™ 4B with bound protein, 4-flow-through off column, 5-final wash, 6-eluted GST fusion protein off column, 7-glutathione sepharose™ 4B bound with thrombin cleaved GST-tag, 8-streptavidin agarose with bound biotinylated thrombin, 9-pure SSI-G1488P protein without GST tag. The predicted size of SSI-G1488P with the GST-tag was 83 kD, and without the GST-tag the predicted size was 56 kD.

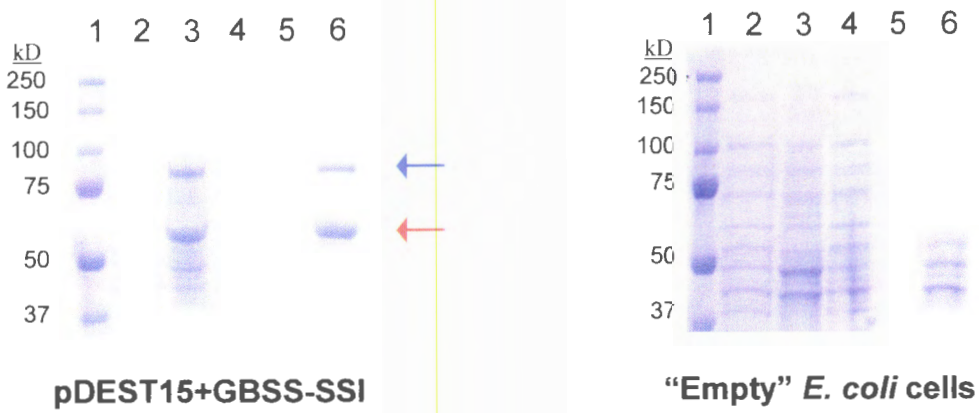


Figure 14: Expression and purification scheme of pDEST15+GBSS-SSI and "empty" *E. coli* cells to determine if GroEL was co-purifying due to affinity to glutathione sepharose™ 4B or its affinity to the chimeric SS enzyme. The blue arrow points to chimeric SS enzyme and red arrow points to GroEL protein. Notice the lack of either band from the "empty" *E. coli* cells. Lane assignments are as follows: 1-molecular weight markers, 2-supernatant after sonication, 3-glutathione sepharose™ 4B with bound protein, 4-flow-through off column, 5-final wash, and 6-eluted GST fusion protein off column.

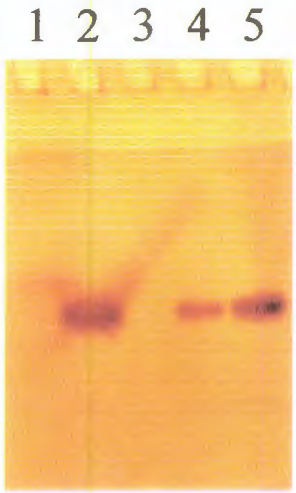
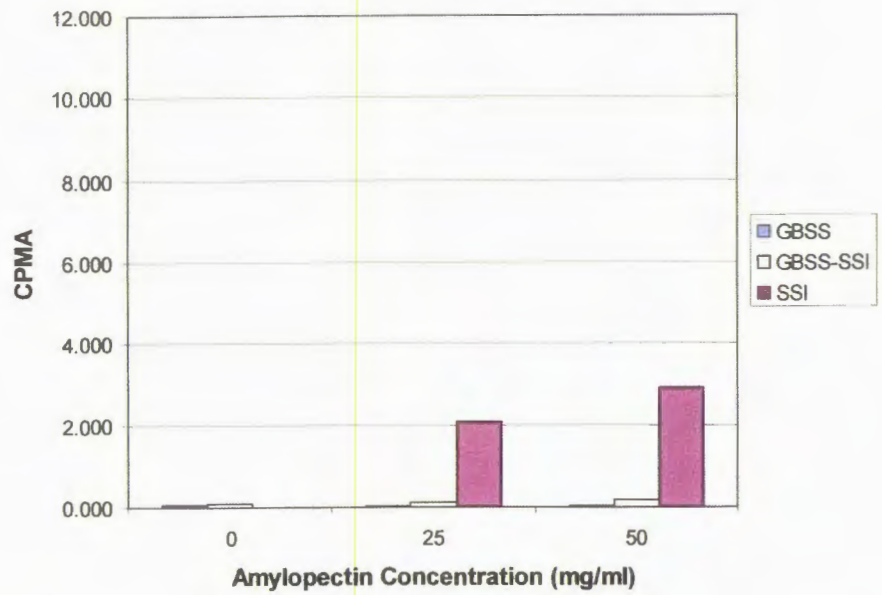


Figure 15: Zymogram analysis of the five purified proteins where a dark band represents the elongation of the rabbit liver glycogen chains in the gel and activity of the enzyme. Lane assignments are as follows: 1-wild-type GBSS, 2-wild-type SSI, 3-GBSS-SSI, 4-SSI-P1445G,R1464A, and 5-SSI-G1488P.

A



B

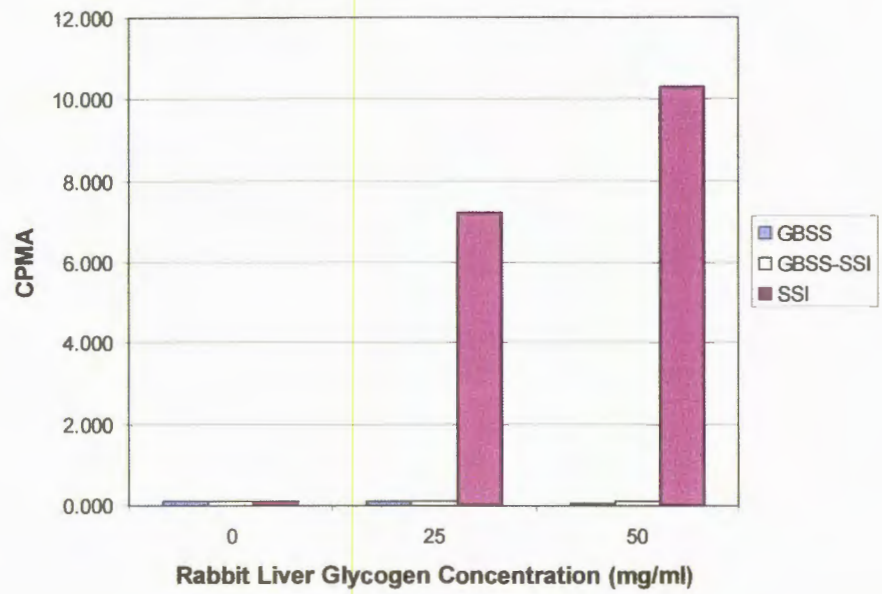


Figure 16: Chart A is a representation of the amount of ADPG (^{14}C) incorporated during an activity assay using amylopectin as a substrate. SSI was the only enzyme that revealed significant activity. Chart B is an illustration of the amount of ADPG (^{14}C) incorporate during an activity assay using rabbit liver glycogen as a substrate. Again, SSI was the only enzyme that exhibited a significant amount of activity.

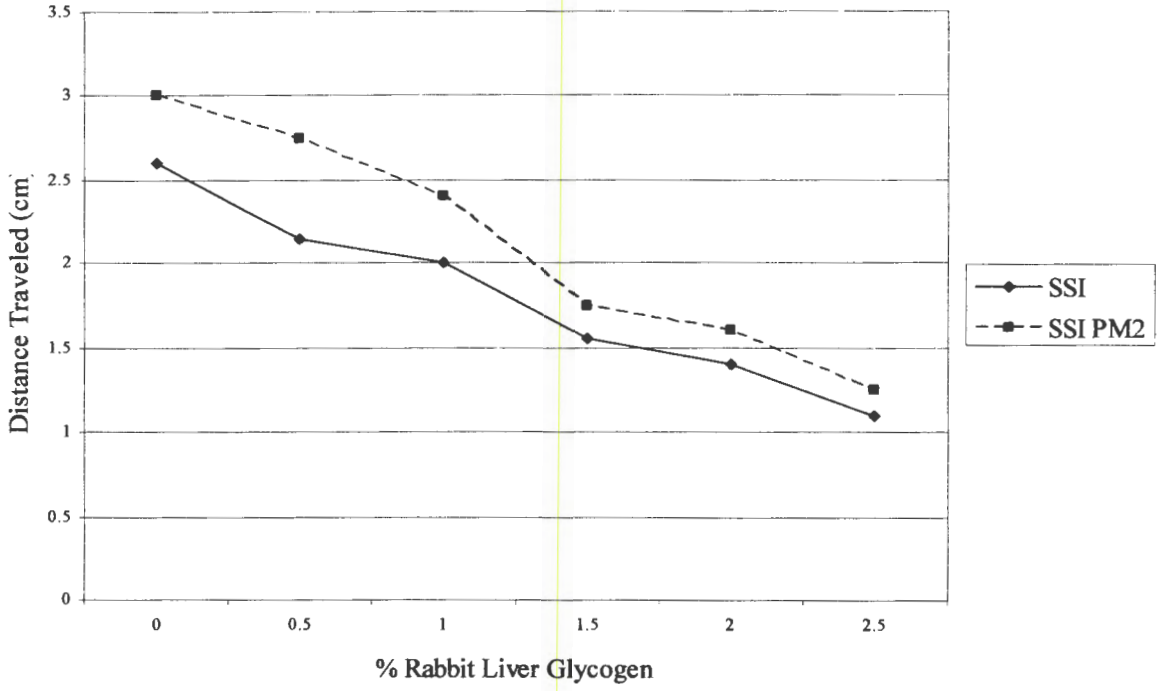


Figure 17: Affinity electrophoresis data from wild-type SSI and SSI-P1445G,R1464A. Both enzymes show a very similar pattern of binding. GBSS, GBSS-SSI, and SSI-G1488P did not reveal any affinity to the glycogen in the gel and ran off.

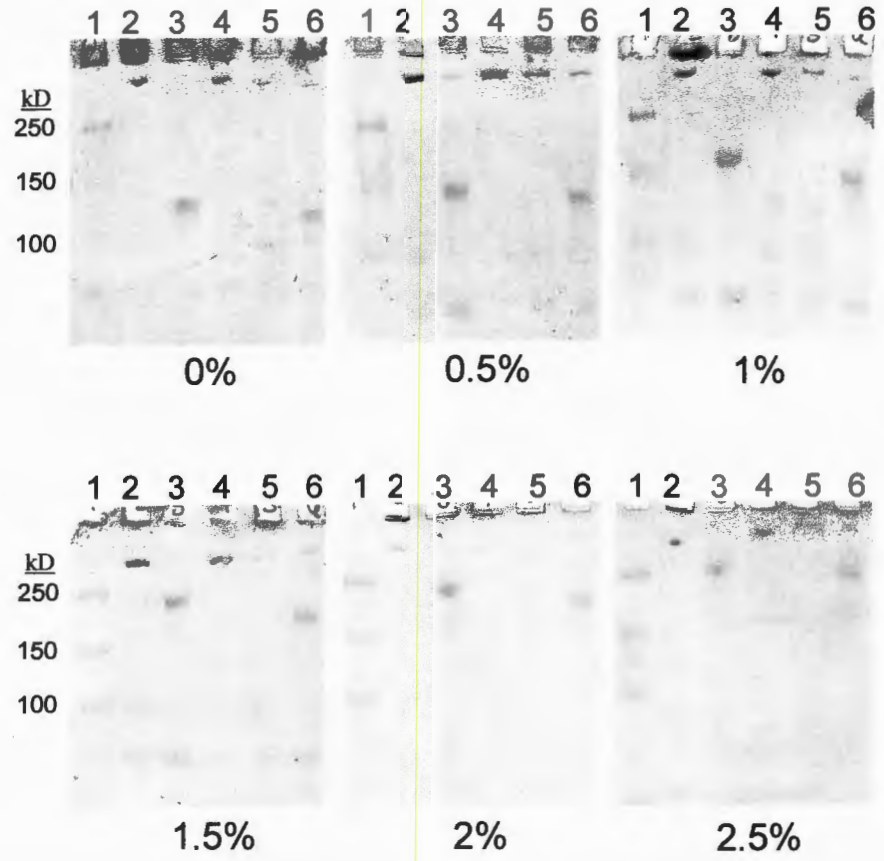


Figure 18: Affinity electrophoresis data from wild-type SSI and SSI-P1445G,R1464A. Both enzymes show a very similar pattern of binding with SSI-P1445G,R1464A showing a slightly lower affinity than SSI. The remaining SS enzymes did not show any affinity to the gel and ran off. Lane assignments are as follows: 1-molecular weight markers, 2-wild-type GBSS, 3-wild-type SSI, 4-GBSS-SSI, 5-SSI-P1445G,R1464A, and 6-SSI-G1488P.

	GBSSI	SSI
Amylose	0.150 ± 0.01	1.06 ± 0.12
Amylopectin	0.054 ± 0.004	0.07 ± 0.03
Glycogen	(no binding)	3.38 ± 0.83

Table 1: Compared K_d values of maize GBSSI and SSI [6].

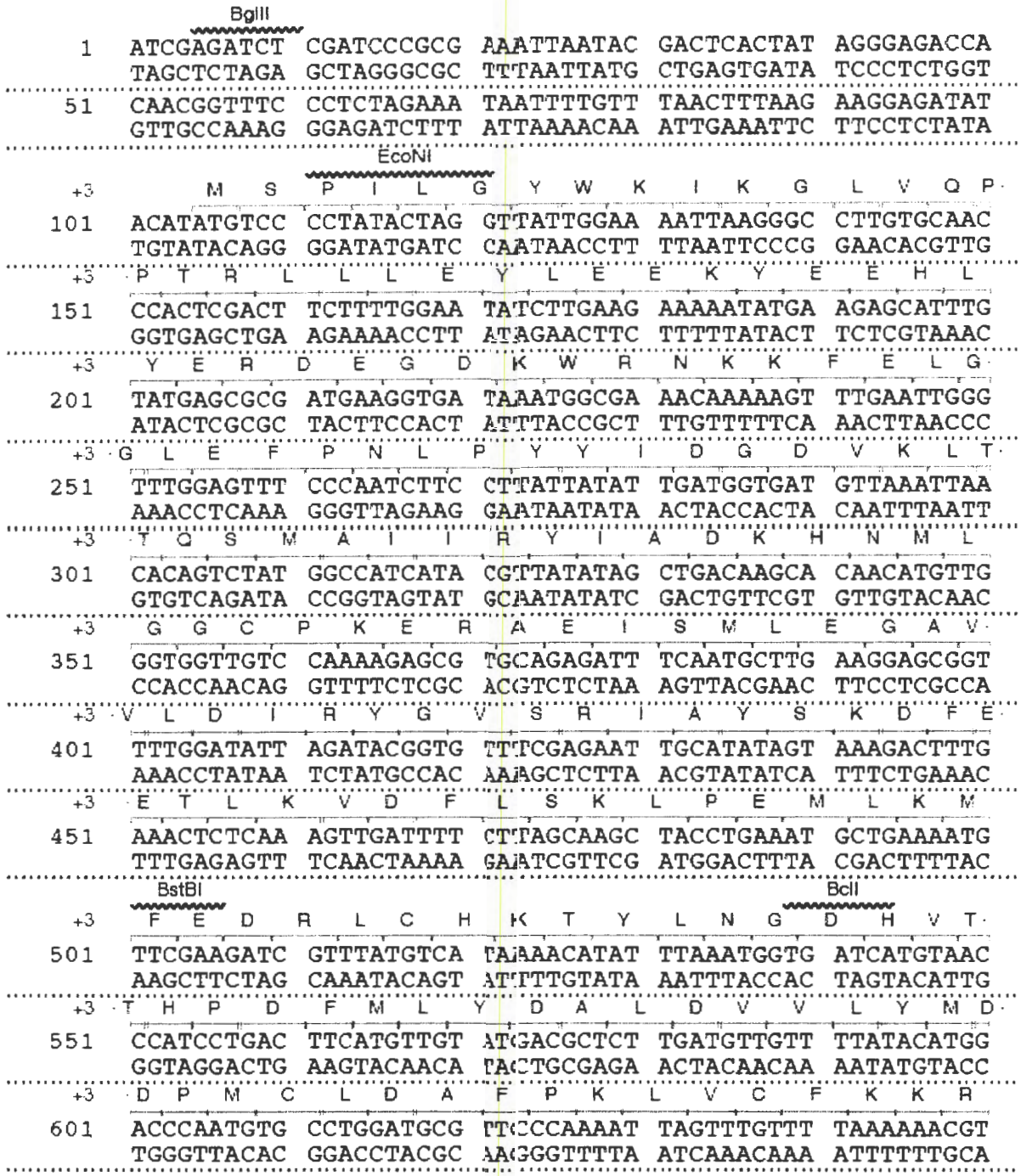
PROTEIN	**THEORETICAL PI**
GBSS	5.49
SSI	5.62
GBSS-SSI	5.18
SSI- P1445G,R1464A	5.70
SSI- G1488P	5.62

Table 2: Expressed and purified proteins' theoretical pI.

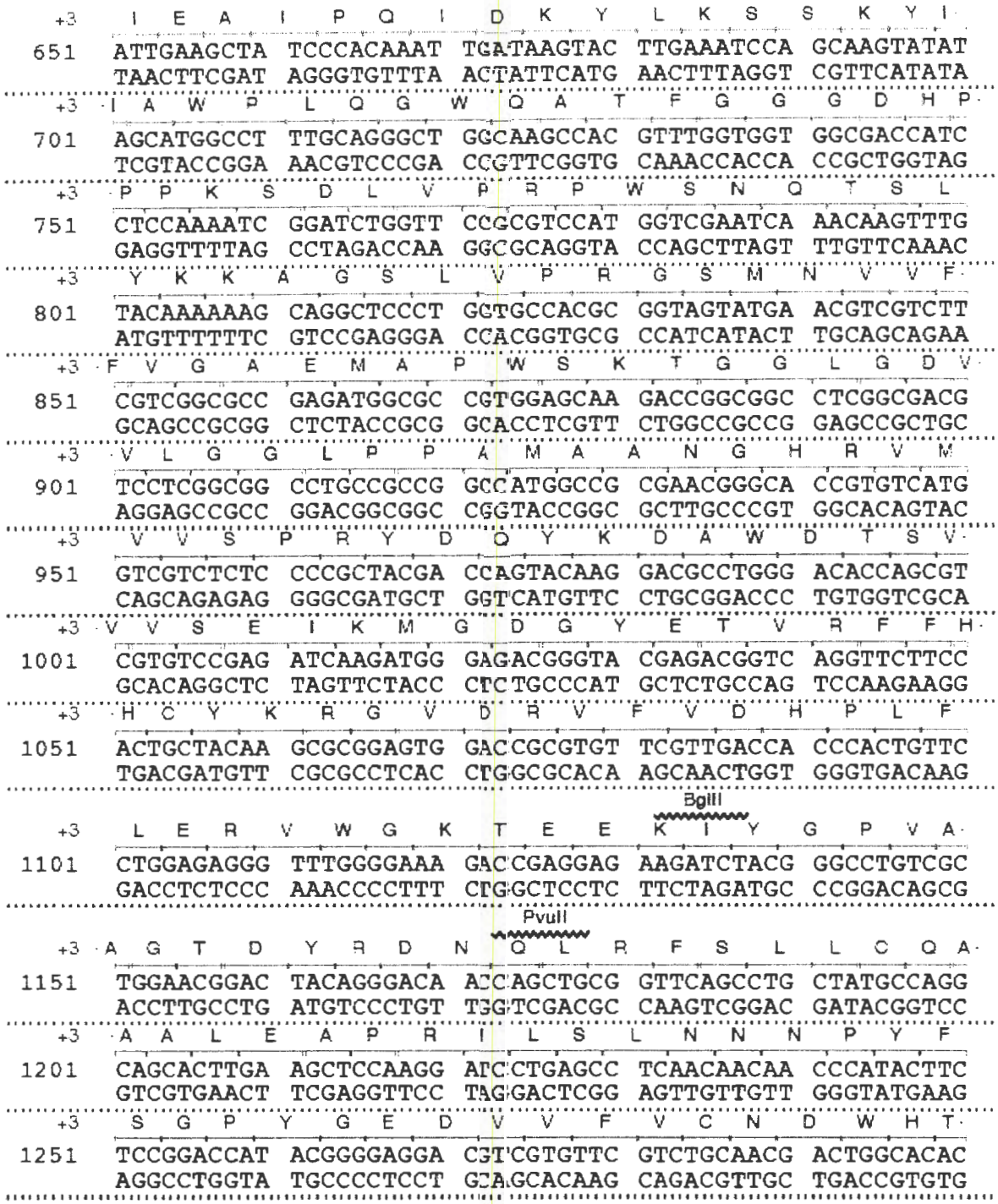
APPENDIX A: PLASMID SEQUENCES

This appendix contains the complete sequences of pDEST15+GBSS, pDEST15+SSI, pDEST15+GBSS-SSI plasmids discussed in Chapter 3. Sequences of pDEST15+SSI-P1445G,R1464A and pDEST15+SSI-G1488P were not included because they are very similar to the pDEST15+SSI sequence. pDEST15+SSI-P1445G,R1464A contains two point mutations: one point mutation changed a C to T transition that would result in a shift from an alanine to an arginine at position 1464 and the second point mutation changed a C to a G that would result in a shift of a proline to a glycine at position 1445. pDEST15+SSI-G1488P contains one point mutation which caused a G to C transversion at position 1488 that would result in a shift from a glycine to a proline. The proteins expressed from these constructs were used to evaluate SS functionality. Sequences and maps (located in Chapter 3) were created using VectorNTI software.

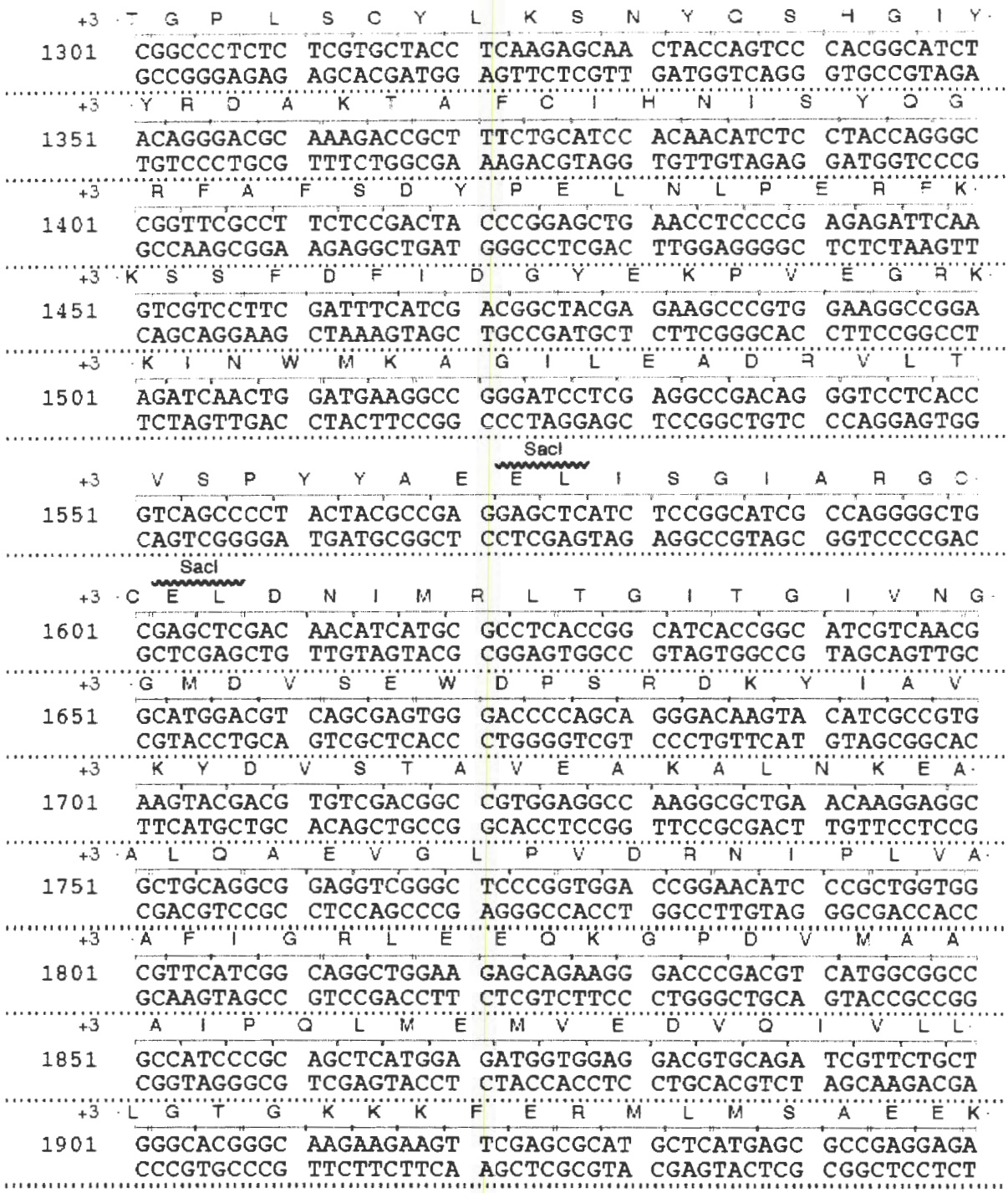
pDEST15+GBSS



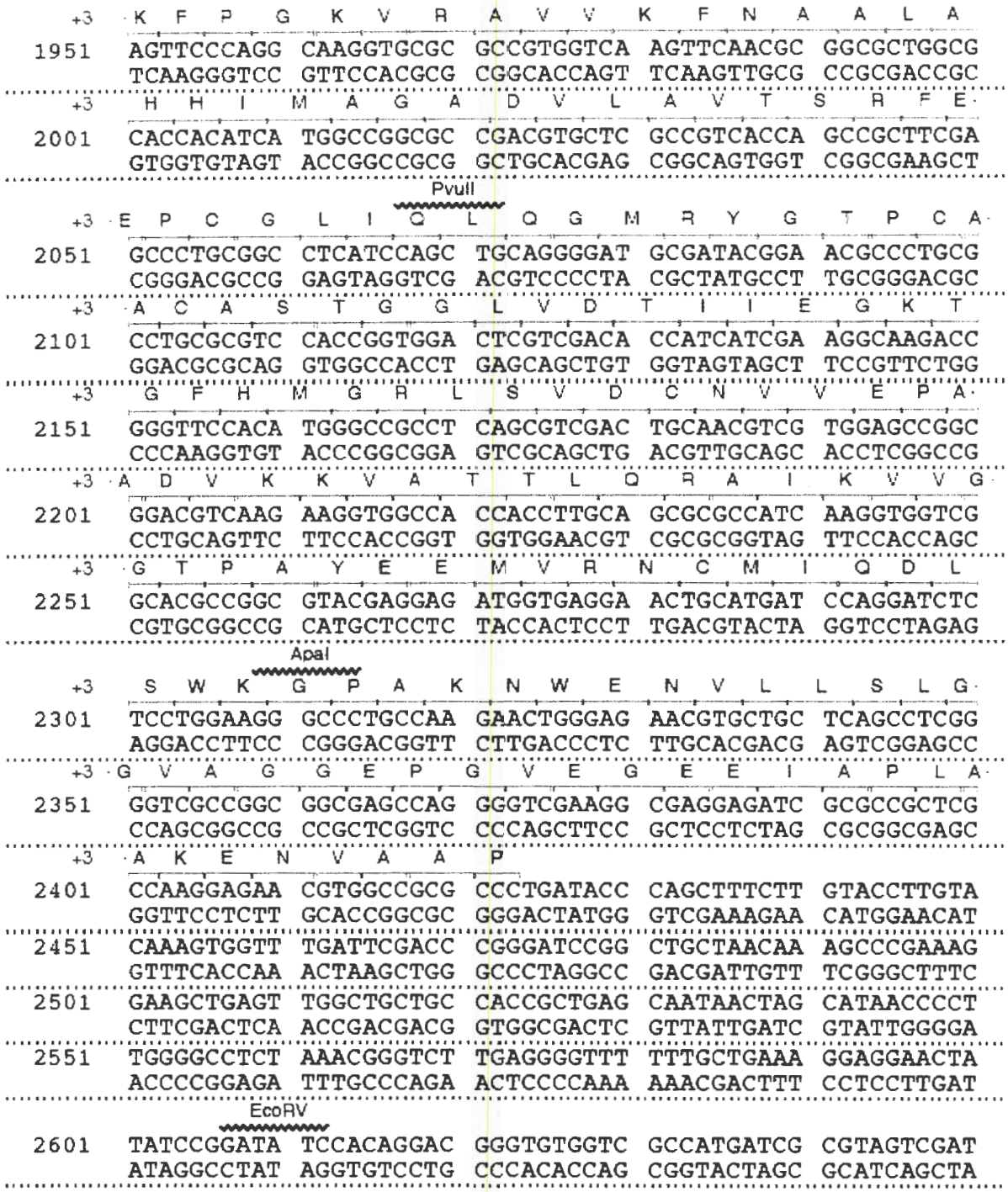
pDEST15+GBSS



pDEST15+GBSS



pDEST15+GBSS



pDEST15+GBSS

```
2651  AGTGGCTCCA AGTAGCGAAG CGAGCAGGAC TGGGCGGCGG CCAAAGCGGT
      TCACCGAGGT TCATCGCTTC GCTCGTCCTG ACCCGCCGCC GGTTTCGCCA

2701  CGGACAGTGC TCCGAGAACG GGTGCGCATA GAAATTGCAT CAACGCATAT
      GCCTGTCACG AGGCTCTTGC CCACGCGTAT CTTTAACGTA GTTGCGTATA

                NheI                                      EcoRV
2751  AGCGCTAGCA GCACGCCATA GTGACTGGCG ATGCTGTCGG AATGGACGAT
      TCGCGATCGT CGTGCGGTAT CACTGACCGC TACGACAGCC TTACCTGCTA

      EcoRV
2801  ATCCCGCAAG AGGCCCGGCA GTACCGGCAT AACCAAGCCT ATGCCTACAG
      TAGGGCGTTC TCCGGGCCGT CATGGCCGTA TTGGTTCGGA TACGGATGTC

2851  CATCCAGGGT GACGGTGCCG AGGATGACGA TGAGCGCATT GTTAGATTTC
      GTAGGTCCCA CTGCCACGGC TCCTACTGCT ACTCGCGTAA CAATCTAAAG

2901  ATACACGGTG CCTGACTGCG TTAGCAATTT AACTGTGATA AACTACCGCA
      TATGTGCCAC GGACTGACGC AATCGTTAAA TTGACACTAT TTGATGGCGT

            HindIII
                    ClaI
2951  TTAAAGCTTA TCGATGATAA GCTGTCAAAC ATGAGAATTC TTGAAGACGA
      AATTTCGAAT AGCTACTATT CGACAGTTTG TACTCTTAAG AACTTCTGCT

3001  AAGGGCCTCG TGATACGCCT ATTTTTATAG GTTAATGTCA TGATAATAAT
      TTCCCGGAGC ACTATGCGGA TAAAAATATC CAATTACAGT ACTATTATTA

3051  GGTTTCTTAG ACGTCAGGTG GCACTTTTCG GGGAAATGTG CGCGGAACCC
      CCAAAGAATC TGCAGTCCAC CGTGAAAAGC CCCTTTACAC GCGCCTTGGG

3101  CTATTTGTTT ATTTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA
      GATAAACAAA TAAAAAGATT TATGTAAGTT TATACATAGG CGAGTACTCT

3151  CAATAACCCT GATAAATGCT TCAATAATAT TGAAAAAGGA AGAGTATGAG
      GTTATTGGGA CTATTTACGA AGTTATTATA ACTTTTTCCT TCTCATACTC

3201  TATTCAACAT TTCCGTGTCG CCCTTATTCC CTTTTTTGCG GCATTTTGCC
      ATAAGTTGTA AAGGCACAGC GGGAATAAGG GAAAAAACGC CGTAAAACGG

3251  TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA AGATGCTGAA
      AAGGACAAAA ACGAGTGGGT CTTTGCGACC ACTTTCATTT TCTACGACTT

                    ApaLI
3301  GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG
      CTAGTCAACC CACGTGCTCA CCCAATGTAG CTTGACCTAG AGTTGTCGCC

3351  TAAGATCCTT GAGAGTTTTC GCCCCGAAGA ACGTTTTCCA ATGATGAGCA
      ATTCTAGGAA CTCTCAAAAG CGGGGCTTCT TGCAAAAGGT TACTACTCGT

3401  CTTTTAAAGT TCTGCTATGT GGCGCGGTAT TATCCCGTGT TGACGCCGGG
      GAAAATTTCA AGACGATACA CCGCGCCATA ATAGGGCACA ACTGCGGCCC

3451  CAAGAGCAAC TCGGTCGCCG CATACACTAT TCTCAGAATG ACTTGGTTGA
      GTTCTCGTTG AGCCAGCGGC GTATGTGATA AGAGTCTTAC TGAACCAACT

3501  GTACTCACCA GTCACAGAAA AGCATCTTAC GGATGGCATG ACAGTAAGAG
      CATGAGTGGT CAGTGTCTTT TCGTAGAATG CCTACCGTAC TGTCATTCTC
```

pDEST15+GBSS

```
3551   AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC GGCCAACTTA
       TTAATACGTC ACGACGGTAT TGGTACTCAC TATTGTGACG CCGGTTGAAT
                  PvuI
3601   CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA
       GAAGACTGTT GCTAGCCTCC TGGCTTCCTC GATTGGCGAA AAAACGTGTT
3651   CATGGGGGAT CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG
       GTACCCCCTA GTACATTGAG CGGAACTAGC AACCCTTGGC CTCGACTTAC
3701   AAGCCATACC AAACGACGAG CGTGACACCA CGATGCCTGC AGCAATGGCA
       TTCGGTATGG TTTGCTGCTC GCACTGTGGT GCTACGGACG TCGTTACCGT
3751   ACAACGTTGC GCAAACTATT AACTGGCGAA CTACTTACTC TAGCTTCCCG
       TGTTGCAACG CGTTTGATAA TTGACCGCTT GATGAATGAG ATCGAAGGGC
3801   GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACTTC
       CGTTGTTAAT TATCTGACCT ACCTCCGCCT ATTTCAACGT CCTGGTGAAG
3851   TGCGCTCGGC CCTTCCGGCT GGCTGGTTTA TTGCTGATAA ATCTGGAGCC
       ACGCGAGCCG GGAAGGCCGA CCGACCAAAT AACGACTATT TAGACCTCGG
3901   GGTGAGCGTG GGTCTCGCGG TATCATTGCA GCACTGGGGC CAGATGGTAA
       CCACTCGCAC CCAGAGCGCC ATAGTAACGT CGTGACCCCG GTCTACCATT
3951   GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG
       CGGGAGGGCA TAGCATCAAT AGATGTGCTG CCCCTCAGTC CGTTGATACC
4001   ATGAACGAAA TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT
       TACTTGCTTT ATCTGTCTAG CGACTCTATC CACGGAGTGA CTAATTCGTA
4051   TGGTAACTGT CAGACCAAGT TTACTCATAT ATACTTTAGA TTGATTTAAA
       ACCATTGACA GTCTGGTTCA AATGAGTATA TATGAAATCT AACTAAATTT
4101   ACTTCATTTT TAATTTAAAA GGATCTAGGT GAAGATCCTT TTTGATAATC
       TGAAGTAAAA ATTAAATTTT CCTAGATCCA CTTCTAGGAA AAACTATTAG
4151   TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG AGCGTCAGAC
       AGTACTGGTT TTAGGGAATT GCACTCAAAA GCAAGGTGAC TCGCAGTCTG
4201   CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT
       GGGCATCTTT TCTAGTTTCC TAGAAGAACT CTAGGAAAAA AAGACGCGCA
4251   AATCTGCTGC TTGCAAACAA AAAAACCACC GCTACCAGCG GTGGTTTGTT
       TTAGACGACG AACGTTTGTT TTTTTGGTGG CGATGGTCGC CACCAAACAA
4301   TGCCGGATCA AGAGCTACCA ACTTCTTTTTC CGAAGGTAAC TGGCTTCAGC
       ACGGCCTAGT TCTCGATGGT TGAAGAAAAAG GCTTCCATTG ACCGAAGTCG
4351   AGAGCGCAGA TACCAAATAC TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA
       TCTCGCGTCT ATGGTTTATG ACAGGAAGAT CACATCGGCA TCAATCCGGT
4401   CCACTTCAAG AACTCTGTAG CACCGCCTAC ATACCTCGCT CTGCTAATCC
       GGTGAAGTTC TTGAGACATC GTGGCGGATG TATGGAGCGA GACGATTAGG
4451   TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT TACCGGGTTG
       ACAATGGTCA CCGACGACGG TCACCGCTAT TCAGCACAGA ATGGCCCAAC
4501   GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG
       CTGAGTTCTG CTATCAATGG CCTATTCCGC GTCGCCAGCC CGACTTGCCC
```

pDEST15+GBSS

ApaLI

```
4551  GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA
      CCCAAGCACG TGTGTCGGGT CGAACCTCGC TTGCTGGATG TGGCTTGACT
4601  GATACCTACA GCGTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA
      CTATGGATGT CGCACTCGAT ACTCTTTCGC GGTGCGAAGG GCTTCCCTCT
4651  AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC
      TTCCGCCTGT CCATAGGCCA TTCGCCGTCC CAGCCTTGTC CTCTCGCGTG
4701  GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT
      CTCCCTCGAA GGTCCCCCTT TGCGGACCAT AGAAATATCA GGACAGCCCA
4751  TTCGCCACCT CTGACTTGAG CGTCGATTTT TGTGATGCTC GTCAGGGGGG
      AAGCGGTGGA GACTGAACTC GCAGCTAAAA ACACTACGAG CAGTCCCCCC
4801  CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTTAC GGTTCCTGGC
      GCCTCGGATA CCTTTTTGCG GTCGTTGCGC CGGAAAAATG CCAAGGACCG
4851  CTTTTGCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT
      GAAAACGACC GGAAAACGAG TGTACAAGAA AGGACGCAAT AGGGGACTAA
4901  CTGTGGATAA CCGTATTACC GCCTTTGAGT GAGCTGATAC CGCTCGCCGC
      GACACCTATT GGCATAATGG CGGAAACTCA CTCGACTATG GCGAGCGGCG
4951  AGCCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG CGGAAGAGCG
      TCGGCTTGCT GGCTCGCGTC GCTCAGTCAC TCGCTCCTTC GCCTTCTCGC
5001  CCTGATGCGG TATTTTCTCC TTACGCATCT GTGCGGTATT TCACACCGCA
      GGACTACGCC ATAAAAGAGG AATGCGTAGA CACGCCATAA AGTGTGGCGT
```

ApaLI

```
5051  TATATGGTGC ACTCTCAGTA CAATCTGCTC TGATGCCGCA TAGTTAAGCC
      ATATACCACG TGAGAGTCAT GTTAGACGAG ACTACGGCGT ATCAATTCGG
5101  AGTATACACT CCGCTATCGC TACGTGACTG GGTCATGGCT GCGCCCCGAC
      TCATATGTGA GGCGATAGCG ATGCACTGAC CCAGTACCGA CGCGGGGCTG
5151  ACCCGCCAAC ACCCGCTGAC GCGCCCTGAC GGGCTTGTCT GCTCCCGGCA
      TGGGCGGTTG TGGGCGACTG CGCGGGACTG CCCGAACAGA CGAGGGCCGT
5201  TCCGCTTACA GACAAGCTGT GACCGTCTCC GGGAGCTGCA TGTGTCAGAG
      AGGCGAATGT CTGTTCGACA CTGGCAGAGG CCCTCGACGT ACACAGTCTC
```

PvuII

```
5251  GTTTTCACCG TCATCACCGA AACGCGCGAG GCAGCTGCGG TAAAGCTCAT
      CAAAAGTGGC AGTAGTGGCT TTGCGCGCTC CGTCGACGCC ATTTCGAGTA
5301  CAGCGTGGTC GTGAAGCGAT TCACAGATGT CTGCCTGTTC ATCCGCGTCC
      GTCGCACCAG CACTTCGCTA AGTGTCTACA GACGGACAAG TAGGCGCAGG
5351  AGCTCGTTGA GTTTCTCCAG AAGCGTTAAT GTCTGGCTTC TGATAAAGCG
      TCGAGCAACT CAAAGAGGTC TTCGCAATTA CAGACCGAAG ACTATTTCGC
5401  GGCCATGTTA AGGGCGGTTT TTTCCTGTTT GGTCACTGAT GCCTCCGTGT
      CCGGTACAAT TCCCGCCAAA AAAGGACAAA CCAGTGACTA CGGAGGCACA
5451  AAGGGGGATT TCTGTTCATG GGGGTAATGA TACCGATGAA ACGAGAGAGG
      TTCCCCCTAA AGACAAGTAC CCCCATTACT ATGGCTACTT TGCTCTCTCC
```

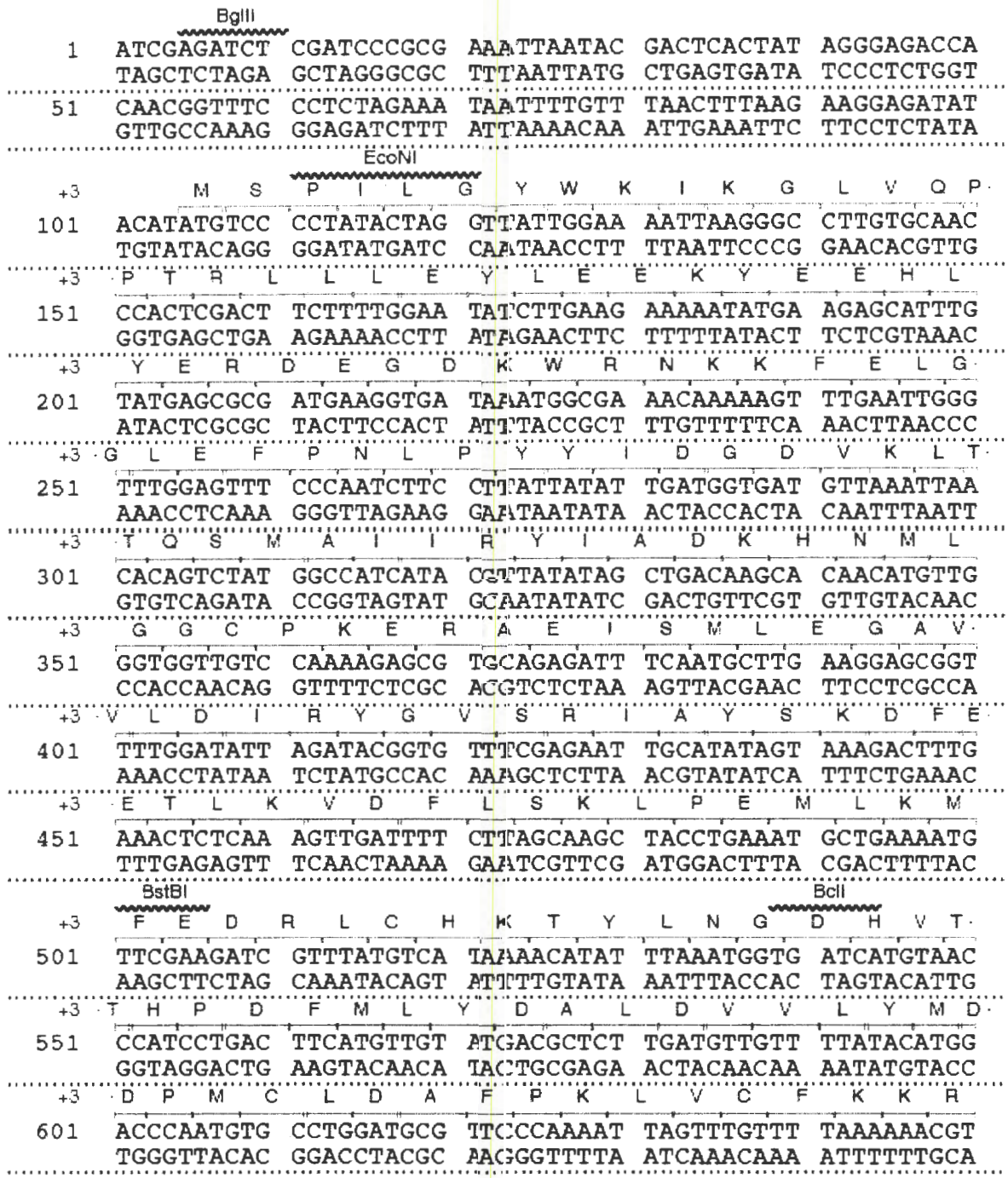
pDEST15+GBSS

```
5501   ATGCTCACGA TACGGGTTAC TGATGATGAA CATGCCCGGT TACTGGAACG
       TACGAGTGCT ATGCCCAATG ACTACTACTT GTACGGGCCA ATGACCTTGC

5551   TTGTGAGGGT AAACAACTGG CGGTATGGAT GCGGCGGGAC CAGAGAAAAA
       AACACTCCCA TTTGTTGACC GCCATACCTA CGCCGCCCTG GTCTCTTTTT

5601   TCACTCAGGG TCAATGCCAG CGCTTCGTTA ATACAGATGT AGGTGTTCCA
       AGTGAGTCCC AGTTACGGTC GCGAAGCAAT TATGTCTACA TCCACAAGGT

5651   CAGGGTAGCC AGCAGCATCC TGCGATGCAG ATCCGGAACA TAATGGTGCA
       GTCCCATCGG TCGTCGTAGG ACGCTACGTC TAGGCCTTGT ATTACCACGT

5701   GGGCGCTGAC TTCCGCGTTT CCAGACTTTA CGAAACACGG AAACCGAAGA
       CCCGCGACTG AAGGCGCAAA GGTCTGAAAT GCTTTGTGCC TTTGGCTTCT

5751   CCATTCATGT TGTTGCTCAG GTCGCAGACG TTTTGCAGCA GCAGTCGCTT
       GGTAAGTACA ACAACGAGTC CAGCGTCTGC AAAACGTCGT CGTCAGCGAA

5801   CACGTTCGCT CGCGTATCGG TGATTCATTC TGCTAACCAG TAAGGCAACC
       GTGCAAGCGA GCGCATAGCC ACTAAGTAAG ACGATTGGTC ATTCCGTTGG

5851   CCGCCAGCCT AGCCGGGTCC TCAACGACAG GAGCACGATC ATGCGCACCC
       GGCGGTCGGA TCGGCCCAGG AGTTGCTGTC CTCGTGCTAG TACGCGTGGG

5901   GTGGCCAGGA CCCAACGCTG CCCGAGATGC GCCGCGTGCG GCTGCTGGAG
       CACCGGTCCT GGGTTGCGAC GGGCTCTACG CGGCGCACGC CGACGACCTC

5951   ATGGCGGACG CGATGGATAT GTTCTGCCAA GGGTTGGTTT GCGCATTCAC
       TACCGCCTGC GCTACCTATA CAAGACGGTT CCCAACCAAA CGCGTAAGTG
                                                  ~~~~~~~~~~
                                                  BsmI

6001   AGTTCTCCGC AAGAATTGAT TGGCTCCAAT TCTTGGAGTG GTGAATCCGT
       TCAAGAGGCG TTCTTAACTA ACCGAGGTTA AGAACCTCAC CACTTAGGCA

6051   TAGCGAGGTG CCGCCGGCTT CCATTCAGGT CGAGGTGGCC CGGCTCCATG
       ATCGCTCCAC GGCGGCCGAA GGTAAGTCCA GCTCCACCGG GCCGAGGTAC

6101   CACCGCGACG CAACGCGGGG AGGCAGACAA GGTATAGGGC GGCGCCTACA
       GTGGCGCTGC GTTGCGCCCC TCCGTCTGTT CCATATCCCG CCGCGGATGT

6151   ATCCATGCCA ACCCGTTCCA TGTGCTCGCC GAGGCGGCAT AAATCGCCGT
       TAGGTACGGT TGGGCAAGGT ACACGAGCGG CTCCGCCGTA TTTAGCGGCA

6201   GACGATCAGC GGTCCAGTGA TCGAAGTTAG GCTGGTAAGA GCCGCGAGCG
       CTGCTAGTCG CCAGGTCACT AGCTTCAATC CGACCATTCT CGGCGCTCGC

6251   ATCCTTGAAG CTGTCCCTGA TGGTCGTCAT CTACCTGCCT GGACAGCATG
       TAGGAACTTC GACAGGGACT ACCAGCAGTA GATGGACGGA CCTGTCGTAC

6301   GCCTGCAACG CGGGCATCCC GATGCCGCCG GAAGCGAGAA GAATCATAAT
       CGGACGTTGC GCCCGTAGGG CTACGGCGGC CTTCGCTCTT CTTAGTATTA
                                  ~~~~~~~~~~
                                  NruI

6351   GGGGAAGGCC ATCCAGCCTC GCGTCGCGAA CGCCAGCAAG ACGTAGCCCA
       CCCCTTCCGG TAGGTCGGAG CGCAGCGCTT GCGGTCGTTC TGCATCGGGT

6401   GCGCGTCGGC CGCCATGCCG GCGATAATGG CCTGCTTCTC GCCGAAACGT
       CGCGCAGCCG GCGGTACGGC CGCTATTACC GGACGAAGAG CGGCTTTGCA

6451   TTGGTGGCGG GACCAGTGAC GAAGGCTTGA GCGAGGGCGT GCAAGATTCC
       AACCACCGCC CTGGTCACTG CTTCCGAACT CGCTCCCGCA CGTTCTAAGG
```

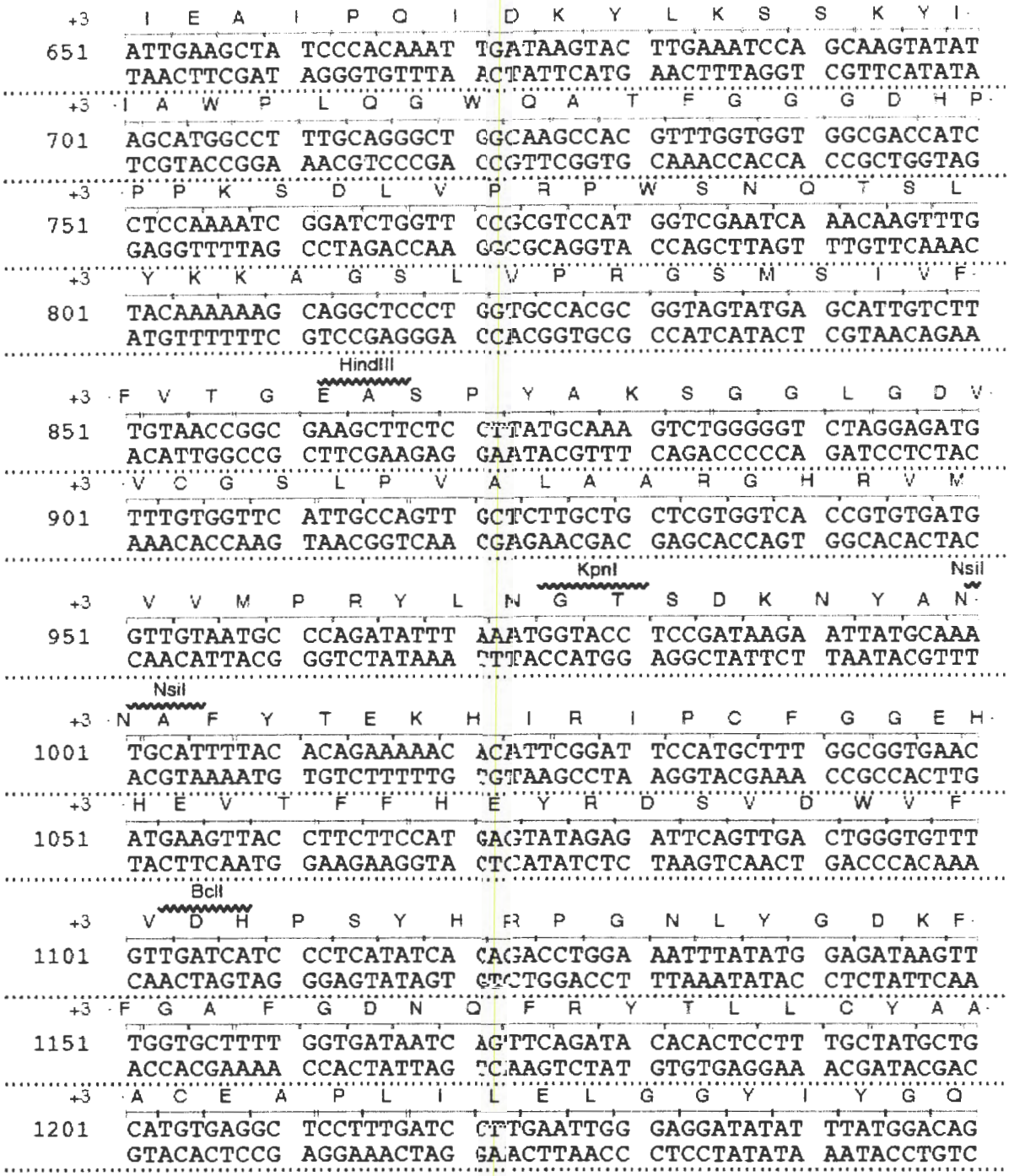
pDEST15+GBSS

```
6501   GAATACCGCA AGCGACAGGC CGATCATCGT CGCGCTCCAG CGAAAGCGGT
       CTTATGGCGT TCGCTGTCCG GCTAGTAGCA GCGCGAGGTC GCTTTCGCCA
6551   CCTCGCCGAA AATGACCCAG AGCGCTGCCG GCACCTGTCC TACGAGTTGC
       GGAGCGGCTT TTACTGGGTC TCGCGACGGC CGTGGACAGG ATGCTCAACG
6601   ATGATAAAGA AGACAGTCAT AAGTGCGGCG ACGATAGTCA TGCCCCGCGC
       TACTATTTCT TCTGTCAGTA TTCACGCCGC TGCTATCAGT ACGGGGCGCG
                                                      PvuI
6651   CCACCGGAAG GAGCTGACTG GGTTGAAGGC TCTCAAGGGC ATCGGTCGAT
       GGTGGCCTTC CTCGACTGAC CCAACTTCCG AGAGTTCCCG TAGCCAGCTA
       PvuI                          EcoNI
6701   CGACGCTCTC CCTTATGCGA CTCCTGCATT AGGAAGCAGC CCAGTAGTAG
       GCTGCGAGAG GGAATACGCT GAGGACGTAA TCCTTCGTCG GGTCATCATC
6751   GTTGAGGCCG TTGAGCACCG CCGCCGCAAG GAATGGTGCA TGCAAGGAGA
       CAACTCCGGC AACTCGTGGC GGCGGCGTTC CTTACCACGT ACGTTCCTCT
6801   TGGCGCCCAA CAGTCCCCCG GCCACGGGGC CTGCCACCAT ACCCACGCCG
       ACCGCGGGTT GTCAGGGGGC CGGTGCCCCG GACGGTGGTA TGGGTGCGGC
6851   AAACAAGCGC TCATGAGCCC GAAGTGGCGA GCCCGATCTT CCCCATCGGT
       TTTGTTCGCG AGTACTCGGG CTTCACCGCT CGGGCTAGAA GGGGTAGCCA
6901   GATGTCGGCG ATATAGGCGC CAGCAACCGC ACCTGTGGCG CCGGTGATGC
       CTACAGCCGC TATATCCGCG GTCGTTGGCG TGGACACCGC GGCCACTACG
6951   CGGCCACGAT GCGTCCGGCG TAGAGG
       GCCGGTGCTA CGCAGGCCGC ATCTCC
```

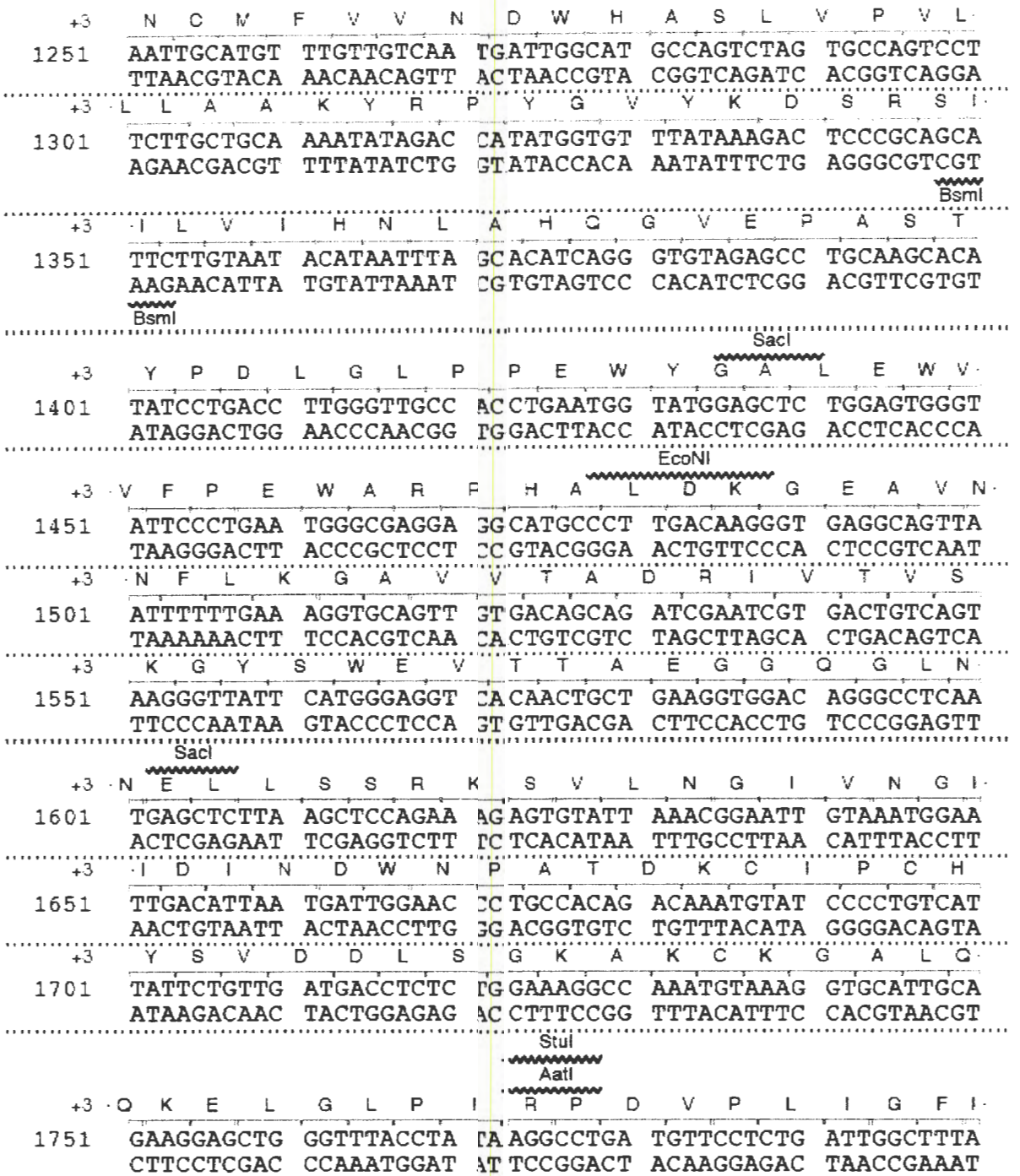
pDEST15+SSI



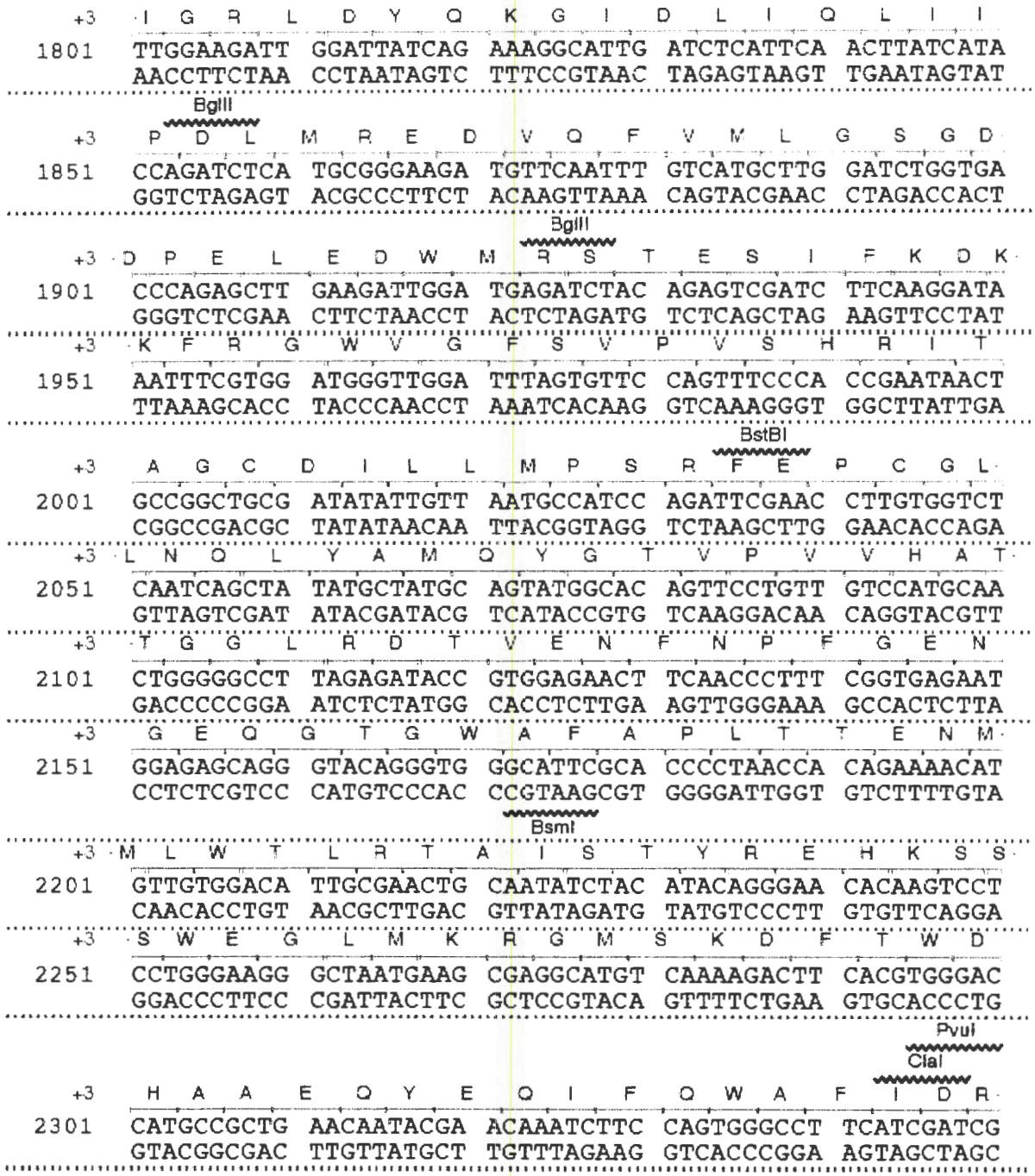
pDEST15+SSI


+3 I E A I P Q I D K Y L K S S K Y I
651 ATTGAAGCTA TCCCACAAAT TGATAAGTAC TTGAAATCCA GCAAGTATAT
TAACTTCGAT AGGGTGTTTA ACTATTCATG AACTTTAGGT CGTTCATATA
+3 I A W P L Q G W Q A T F G G G D H P
701 AGCATGGCCT TTGCAGGGCT GGCAAGCCAC GTTTGGTGGT GGCGACCATC
TCGTACCGGA AACGTCCCGA CCGTTCGGTG CAAACCACCA CCGCTGGTAG
+3 P P K S D L V P R P W S N Q T S L
751 CTCCAAAATC GGATCTGGTT CCGCGTCCAT GGTCGAATCA AACAAGTTTG
GAGGTTTTAG CCTAGACCAA GGCGCAGGTA CCAGCTTAGT TTGTTCAAAC
+3 Y K K A G S L V P R G S M S I V F
801 TACAAAAAAG CAGGCTCCCT GGTGCCACGC GGTAGTATGA GCATTGTCTT
ATGTTTTTTC GTCCGAGGGA CCACGGTGCG CCATCATACT CGTAACAGAA
HindIII
+3 F V T G E A S P Y A K S G G L G D V
851 TGTAACCGGC GAAGCTTCTC CTTATGCAAA GTCTGGGGGT CTAGGAGATG
ACATTGGCCG CTTCGAAGAG GAATACGTTT CAGACCCCCA GATCCTCTAC
+3 V C G S L P V A L A A R G H R V M
901 TTTGTGGTTC ATTGCCAGTT GCTCTTGCTG CTCGTGGTCA CCGTGTGATG
AAACACCAAG TAACGGTCAA CGAGAACGAC GAGCACCAGT GGCACACTAC
KpnI NsiI
+3 V V M P R Y L N G T S D K N Y A N
951 GTTGTAATGC CCAGATATTT AAATGGTACC TCCGATAAGA ATTATGCAAA
CAACATTACG GGTCTATAAA TTTACCATGG AGGCTATTCT TAATACGTTT
NsiI
+3 N A F Y T E K H I R I P C F G G E H
1001 TGCATTTTAC ACAGAAAAAC ACATTCGGAT TCCATGCTTT GGCGGTGAAC
ACGTAAAATG TGTCTTTTTG TGTAAGCCTA AGGTACGAAA CCGCCACTTG
+3 H E V T F F H E Y R D S V D W V F
1051 ATGAAGTTAC CTTCTTCCAT GAGTATAGAG ATTCAGTTGA CTGGGTGTTT
TACTTCAATG GAAGAAGGTA CTCATATCTC TAAGTCAACT GACCCACAAA
BclI
+3 V D H P S Y H R P G N L Y G D K F
1101 GTTGATCATC CCTCATATCA CAGACCTGGA AATTTATATG GAGATAAGTT
CAACTAGTAG GGAGTATAGT GTCTGGACCT TTAAATATAC CTCTATTCAA
+3 F G A F G D N Q F R Y T L L C Y A A
1151 TGGTGCTTTT GGTGATAATC AGTTCAGATA CACACTCCTT TGCTATGCTG
ACCACGAAAA CCACTATTAG TCAAGTCTAT GTGTGAGGAA ACGATACGAC
+3 A C E A P L I L E L G G Y I Y G Q
1201 CATGTGAGGC TCCTTTGATC CTTGAATTGG GAGGATATAT TTATGGACAG
GTACACTCCG AGGAAACTAG GAACTTAACC CTCCTATATA AATACCTGTC

pDEST15+SSI


+3 N C M F V V N D W H A S L V P V L
1251 AATTGCATGT TTGTTGTCAA TGATTGGCAT GCCAGTCTAG TGCCAGTCCT
TTAACGTACA AACAACAGTT ACTAACCGTA CGGTCAGATC ACGGTCAGGA
+3 L L A A K Y R P Y G V Y K D S R S I
1301 TCTTGCTGCA AAATATAGAC CATATGGTGT TTATAAAGAC TCCCGCAGCA
AGAACGACGT TTTATATCTG GTATACCACA AATATTTCTG AGGGCGTCGT
BsmI
+3 I L V I H N L A H Q G V E P A S T
1351 TTCTTGTAAT ACATAATTTA GCACATCAGG GTGTAGAGCC TGCAAGCACA
AAGAACATTA TGTATTAAAT CGTGTAGTCC CACATCTCGG ACGTTCGTGT
BsmI
SacI
+3 Y P D L G L P P E W Y G A L E W V
1401 TATCCTGACC TTGGGTTGCC ACCTGAATGG TATGGAGCTC TGGAGTGGGT
ATAGGACTGG AACCCAACGG TGGACTTACC ATACCTCGAG ACCTCACCCA
EcoNI
+3 V F P E W A R P H A L D K G E A V N
1451 ATTCCCTGAA TGGGCGAGGA GGCATGCCCT TGACAAGGGT GAGGCAGTTA
TAAGGGACTT ACCCGCTCCT CCGTACGGGA ACTGTTCCCA CTCCGTCAAT
+3 N F L K G A V V T A D R I V T V S
1501 ATTTTTTGAA AGGTGCAGTT GTGACAGCAG ATCGAATCGT GACTGTCAGT
TAAAAAACTT TCCACGTCAA CACTGTCGTC TAGCTTAGCA CTGACAGTCA
+3 K G Y S W E V T T A E G G Q G L N
1551 AAGGGTTATT CATGGGAGGT CACAACTGCT GAAGGTGGAC AGGGCCTCAA
TTCCCAATAA GTACCCTCCA GTGTTGACGA CTTCCACCTG TCCCGGAGTT
SacI
+3 N E L L S S R K S V L N G I V N G I
1601 TGAGCTCTTA AGCTCCAGAA AGAGTGTATT AAACGGAATT GTAAATGGAA
ACTCGAGAAT TCGAGGTCTT TCTCACATAA TTTGCCTTAA CATTTACCTT
+3 I D I N D W N P A T D K C I P C H
1651 TTGACATTAA TGATTGGAAC CCTGCCACAG ACAAATGTAT CCCCTGTCAT
AACTGTAATT ACTAACCTTG GGACGGTGTC TGTTTACATA GGGGACAGTA
+3 Y S V D D L S G K A K C K G A L Q
1701 TATTCTGTTG ATGACCTCTC TGGAAAGGCC AAATGTAAAG GTGCATTGCA
ATAAGACAAC TACTGGAGAG ACCTTTCCGG TTTACATTTC CACGTAACGT
StuI
AatI
+3 Q K E L G L P I R P D V P L I G F I
1751 GAAGGAGCTG GGTTTACCTA TAAGGCCTGA TGTTCCTCTG ATTGGCTTTA
CTTCCTCGAC CCAAATGGAT ATTCCGGACT ACAAGGAGAC TAACCGAAAT

pDEST15+SSI


+3 I G R L D Y Q K G I D L I Q L I I
1801 TTGGAAGATT GGATTATCAG AAAGGCATTG ATCTCATTCA ACTTATCATA
AACCTTCTAA CCTAATAGTC TTTCCGTAAC TAGAGTAAGT TGAATAGTAT
BglII
+3 P D L M R E D V Q F V M L G S G D
1851 CCAGATCTCA TGCGGGAAGA TGTTCAATTT GTCATGCTTG GATCTGGTGA
GGTCTAGAGT ACGCCCTTCT ACAAGTTAAA CAGTACGAAC CTAGACCACT
BglII
+3 D P E L E D W M R S T E S I F K D K
1901 CCCAGAGCTT GAAGATTGGA TGAGATCTAC AGAGTCGATC TTCAAGGATA
GGGTCTCGAA CTTCTAACCT ACTCTAGATG TCTCAGCTAG AAGTTCCTAT
+3 K F R G W V G F S V P V S H R I T
1951 AATTTCGTGG ATGGGTTGGA TTTAGTGTTC CAGTTTCCCA CCGAATAACT
TTAAAGCACC TACCCAACCT AAATCACAAG GTCAAAGGGT GGCTTATTGA
BstBI
+3 A G C D I L L M P S R F E P C G L
2001 GCCGGCTGCG ATATATTGTT AATGCCATCC AGATTCGAAC CTTGTGGTCT
CGGCCGACGC TATATAACAA TTACGGTAGG TCTAAGCTTG GAACACCAGA
+3 L N Q L Y A M Q Y G T V P V V H A T
2051 CAATCAGCTA TATGCTATGC AGTATGGCAC AGTTCCTGTT GTCCATGCAA
GTTAGTCGAT ATACGATACG TCATACCGTG TCAAGGACAA CAGGTACGTT
+3 T G G L R D T V E N F N P F G E N
2101 CTGGGGGCCT TAGAGATACC GTGGAGAACT TCAACCCTTT CGGTGAGAAT
GACCCCCGGA ATCTCTATGG CACCTCTTGA AGTTGGGAAA GCCACTCTTA
+3 G E Q G T G W A F A P L T T E N M
2151 GGAGAGCAGG GTACAGGGTG GGCATTCGCA CCCCTAACCA CAGAAAACAT
CCTCTCGTCC CATGTCCCAC CCGTAAGCGT GGGGATTGGT GTCTTTTGTA
BsmI
+3 M L W T L R T A I S T Y R E H K S S
2201 GTTGTGGACA TTGCGAACTG CAATATCTAC ATACAGGGAA CACAAGTCCT
CAACACCTGT AACGCTTGAC GTTATAGATG TATGTCCCTT GTGTTCAGGA
+3 S W E G L M K R G M S K D F T W D
2251 CCTGGGAAGG GCTAATGAAG CGAGGCATGT CAAAAGACTT CACGTGGGAC
GGACCCTTCC CGATTACTTC GCTCCGTACA GTTTTCTGAA GTGCACCCTG
PvuI
ClaI
+3 H A A E Q Y E Q I F Q W A F I D R
2301 CATGCCGCTG AACAATACGA ACAAATCTTC CAGTGGGCCT TCATCGATCG
GTACGGCGAC TTGTTATGCT TGTTTAGAAG GTCACCCGGA AGTAGCTAGC

pDEST15+SSI

```
      +3  R  P  Y  V  M  *
2351  ACCCTATGTC ATGTAAGACC CAGCTTTCTT GTACTTGTAC AAAGTGGTTT
      TGGGATACAG TACATTCTGG GTCGAAAGAA CATGAACATG TTTCACCAAA
                AvaI
2401  GATTCGACCC GGGATCCGGC TGCTAACAAA GCCCGAAAGG AAGCTGAGTT
      CTAAGCTGGG CCCTAGGCCG ACGATTGTTT CGGGCTTTCC TTCGACTCAA
2451  GGCTGCTGCC ACCGCTGAGC AATAACTAGC ATAACCCCTT GGGGCCTCTA
      CCGACGACGG TGGCGACTCG TTATTGATCG TATTGGGGAA CCCCGGAGAT
                                                  EcoRV
2501  AACGGGTCTT GAGGGGTTTT TTGCTGAAAG GAGGAACTAT ATCCGGATAT
      TTGCCCAGAA CTCCCCAAAA AACGACTTTC CTCCTTGATA TAGGCCTATA
      EcoRV
2551  CCACAGGACG GGTGTGGTCG CCATGATCGC GTAGTCGATA GTGGCTCCAA
      GGTGTCCTGC CCACACCAGC GGTACTAGCG CATCAGCTAT CACCGAGGTT
2601  GTAGCGAAGC GAGCAGGACT GGGCGGCGGC CAAAGCGGTC GGACAGTGCT
      CATCGCTTCG CTCGTCCTGA CCCGCCGCCG GTTTCGCCAG CCTGTCACGA
                                                  NheI
2651  CCGAGAACGG GTGCGCATAG AAATTGCATC AACGCATATA GCGCTAGCAG
      GGCTCTTGCC CACGCGTATC TTTAACGTAG TTGCGTATAT CGCGATCGTC
                                        EcoRV
2701  CACGCCATAG TGACTGGCGA TGCTGTCGGA ATGGACGATA TCCCGCAAGA
      GTGCGGTATC ACTGACCGCT ACGACAGCCT TACCTGCTAT AGGGCGTTCT
2751  GGCCCGGCAG TACCGGCATA ACCAAGCCTA TGCCTACAGC ATCCAGGGTG
      CCGGGCCGTC ATGGCCGTAT TGGTTCGGAT ACGGATGTCG TAGGTCCCAC
2801  ACGGTGCCGA GGATGACGAT GAGCGCATTG TTAGATTTCA TACACGGTGC
      TGCCACGGCT CCTACTGCTA CTCGCGTAAC AATCTAAAGT ATGTGCCACG
                                                  HindIII
                                                      ClaI
2851  CTGACTGCGT TAGCAATTTA ACTGTGATAA ACTACCGCAT TAAAGCTTAT
      GACTGACGCA ATCGTTAAAT TGACACTATT TGATGGCGTA ATTTCGAATA
      ClaI
2901  CGATGATAAG CTGTCAAACA TGAGAATTCT TGAAGACGAA AGGGCCTCGT
      GCTACTATTC GACAGTTTGT ACTCTTAAGA ACTTCTGCTT TCCCGGAGCA
                                        BspHI
2951  GATACGCCTA TTTTTATAGG TTAATGTCAT GATAATAATG GTTTCTTAGA
      CTATGCGGAT AAAAATATCC AATTACAGTA CTATTATTAC CAAAGAATCT
3001  CGTCAGGTGG CACTTTTCGG GGAAATGTGC GCGGAACCCC TATTTGTTTA
      GCAGTCCACC GTGAAAAGCC CCTTTACACG CGCCTTGGGG ATAAACAAAT
                                        BspHI
3051  TTTTTCTAAA TACATTCAAA TATGTATCCG CTCATGAGAC AATAACCCTG
      AAAAAGATTT ATGTAAGTTT ATACATAGGC GAGTACTCTG TTATTGGGAC
```

pDEST15+SSI

```
3101   ATAAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCAACATT
       TATTTACGAA GTTATTATAA CTTTTTCCTT CTCATACTCA TAAGTTGTAA
3151   TCCGTGTCGC CCTTATTCCC TTTTTTGCGG CATTTTGCCT TCCTGTTTTT
       AGGCACAGCG GGAATAAGGG AAAAAACGCC GTAAAACGGA AGGACAAAAA
                                                      ApaLI
3201   GCTCACCCAG AAACGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG
       CGAGTGGGTC TTTGCGACCA CTTTCATTTT CTACGACTTC TAGTCAACCC
         ApaLI
3251   TGCACGAGTG GGTTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG
       ACGTGCTCAC CCAATGTAGC TTGACCTAGA GTTGTCGCCA TTCTAGGAAC
3301   AGAGTTTTCG CCCCGAAGAA CGTTTTCCAA TGATGAGCAC TTTTAAAGTT
       TCTCAAAAGC GGGGCTTCTT GCAAAAGGTT ACTACTCGTG AAAATTTCAA
3351   CTGCTATGTG GCGCGGTATT ATCCCGTGTT GACGCCGGGC AAGAGCAACT
       GACGATACAC CGCGCCATAA TAGGGCACAA CTGCGGCCCG TTCTCGTTGA
3401   CGGTCGCCGC ATACACTATT CTCAGAATGA CTTGGTTGAG TACTCACCAG
       GCCAGCGGCG TATGTGATAA GAGTCTTACT GAACCAACTC ATGAGTGGTC
3451   TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCAGT
       AGTGTCTTTT CGTAGAATGC CTACCGTACT GTCATTCTCT TAATACGTCA
                                                      PvuI
3501   GCTGCCATAA CCATGAGTGA TAACACTGCG GCCAACTTAC TTCTGACAAC
       CGACGGTATT GGTACTCACT ATTGTGACGC CGGTTGAATG AAGACTGTTG
         PvuI
3551   GATCGGAGGA CCGAAGGAGC TAACCGCTTT TTTGCACAAC ATGGGGGATC
       CTAGCCTCCT GGCTTCCTCG ATTGGCGAAA AAACGTGTTG TACCCCCTAG
3601   ATGTAACTCG CCTTGATCGT TGGGAACCGG AGCTGAATGA AGCCATACCA
       TACATTGAGC GGAACTAGCA ACCCTTGGCC TCGACTTACT TCGGTATGGT
3651   AACGACGAGC GTGACACCAC GATGCCTGCA GCAATGGCAA CAACGTTGCG
       TTGCTGCTCG CACTGTGGTG CTACGGACGT CGTTACCGTT GTTGCAACGC
3701   CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA
       GTTTGATAAT TGACCGCTTG ATGAATGAGA TCGAAGGGCC GTTGTTAATT
3751   TAGACTGGAT GGAGGCGGAT AAAGTTGCAG GACCACTTCT GCGCTCGGCC
       ATCTGACCTA CCTCCGCCTA TTTCAACGTC CTGGTGAAGA CGCGAGCCGG
3801   CTTCCGGCTG GCTGGTTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG
       GAAGGCCGAC CGACCAAATA ACGACTATTT AGACCTCGGC CACTCGCACC
3851   GTCTCGCGGT ATCATTGCAG CACTGGGGCC AGATGGTAAG CCCTCCCGTA
       CAGAGCGCCA TAGTAACGTC GTGACCCCGG TCTACCATTC GGGAGGGCAT
3901   TCGTAGTTAT CTACACGACG GGGAGTCAGG CAACTATGGA TGAACGAAAT
       AGCATCAATA GATGTGCTGC CCCTCAGTCC GTTGATACCT ACTTGCTTTA
3951   AGACAGATCG CTGAGATAGG TGCCTCACTG ATTAAGCATT GGTAACTGTC
       TCTGTCTAGC GACTCTATCC ACGGAGTGAC TAATTCGTAA CCATTGACAG
4001   AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAAA CTTCATTTTT
       TCTGGTTCAA ATGAGTATAT ATGAAATCTA ACTAAATTTT GAAGTAAAAA
```

pDEST15+SSI

```
                                                                  BspHI
4051  AATTTAAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA
      TTAAATTTTC CTAGATCCAC TTCTAGGAAA AACTATTAGA GTACTGGTTT
4101  ATCCCTTAAC GTGAGTTTTC GTTCCACTGA GCGTCAGACC CCGTAGAAAA
      TAGGGAATTG CACTCAAAAG CAAGGTGACT CGCAGTCTGG GGCATCTTTT
4151  GATCAAAGGA TCTTCTTGAG ATCCTTTTTT TCTGCGCGTA ATCTGCTGCT
      CTAGTTTCCT AGAAGAACTC TAGGAAAAAA AGACGCGCAT TAGACGACGA
4201  TGCAAACAAA AAAACCACCG CTACCAGCGG TGGTTTGTTT GCCGGATCAA
      ACGTTTGTTT TTTTGGTGGC GATGGTCGCC ACCAAACAAA CGGCCTAGTT
4251  GAGCTACCAA CTCTTTTTCC GAAGGTAACT GGCTTCAGCA GAGCGCAGAT
      CTCGATGGTT GAGAAAAAGG CTTCCATTGA CCGAAGTCGT CTCGCGTCTA
4301  ACCAAATACT GTCCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA
      TGGTTTATGA CAGGAAGATC ACATCGGCAT CAATCCGGTG GTGAAGTTCT
4351  ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAATCCT GTTACCAGTG
      TGAGACATCG TGGCGGATGT ATGGAGCGAG ACGATTAGGA CAATGGTCAC
4401  GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG
      CGACGACGGT CACCGCTATT CAGCACAGAA TGGCCCAACC TGAGTTCTGC
                                                          ApaLI
4451  ATAGTTACCG GATAAGGCGC AGCGGTCGGG CTGAACGGGG GGTTCGTGCA
      TATCAATGGC CTATTCCGCG TCGCCAGCCC GACTTGCCCC CCAAGCACGT
      ApaLI
4501  CACAGCCCAG CTTGGAGCGA ACGACCTACA CCGAACTGAG ATACCTACAG
      GTGTCGGGTC GAACCTCGCT TGCTGGATGT GGCTTGACTC TATGGATGTC
4551  CGTGAGCTAT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGGCGGACAG
      GCACTCGATA CTCTTTCGCG GTGCGAAGGG CTTCCCTCTT TCCGCCTGTC
4601  GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCACG AGGGAGCTTC
      CATAGGCCAT TCGCCGTCCC AGCCTTGTCC TCTCGCGTGC TCCCTCGAAG
4651  CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGGTT TCGCCACCTC
      GTCCCCCTTT GCGGACCATA GAAATATCAG GACAGCCCAA AGCGGTGGAG
4701  TGACTTGAGC GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG
      ACTGAACTCG CAGCTAAAAA CACTACGAGC AGTCCCCCCG CCTCGGATAC
4751  GAAAAACGCC AGCAACGCGG CCTTTTTACG GTTCCTGGCC TTTTGCTGGC
      CTTTTTGCGG TCGTTGCGCC GGAAAAATGC CAAGGACCGG AAAACGACCG
4801  CTTTTGCTCA CATGTTCTTT CCTGCGTTAT CCCCTGATTC TGTGGATAAC
      GAAAACGAGT GTACAAGAAA GGACGCAATA GGGGACTAAG ACACCTATTG
4851  CGTATTACCG CCTTTGAGTG AGCTGATACC GCTCGCCGCA GCCGAACGAC
      GCATAATGGC GGAAACTCAC TCGACTATGG CGAGCGGCGT CGGCTTGCTG
4901  CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC GGAAGAGCGC CTGATGCGGT
      GCTCGCGTCG CTCAGTCACT CGCTCCTTCG CCTTCTCGCG GACTACGCCA
                                                          ApaLI
4951  ATTTTCTCCT TACGCATCTG TGCGGTATTT CACACCGCAT ATATGGTGCA
      TAAAAGAGGA ATGCGTAGAC ACGCCATAAA GTGTGGCGTA TATACCACGT
```

pDEST15+SSI

```
          ApaLI
5001      CTCTCAGTAC AATCTGCTCT GATGCCGCAT AGTTAAGCCA GTATACACTC
          GAGAGTCATG TTAGACGAGA CTACGGCGTA TCAATTCGGT CATATGTGAG
5051      CGCTATCGCT ACGTGACTGG GTCATGGCTG CGCCCCGACA CCCGCCAACA
          GCGATAGCGA TGCACTGACC CAGTACCGAC GCGGGGCTGT GGGCGGTTGT
5101      CCCGCTGACG CGCCCTGACG GGCTTGTCTG CTCCCGGCAT CCGCTTACAG
          GGGCGACTGC GCGGGACTGC CCGAACAGAC GAGGGCCGTA GGCGAATGTC
5151      ACAAGCTGTG ACCGTCTCCG GGAGCTGCAT GTGTCAGAGG TTTTCACCGT
          TGTTCGACAC TGGCAGAGGC CCTCGACGTA CACAGTCTCC AAAAGTGGCA
                                    PvuII
5201      CATCACCGAA ACGCGCGAGG CAGCTGCGGT AAAGCTCATC AGCGTGGTCG
          GTAGTGGCTT TGCGCGCTCC GTCGACGCCA TTTCGAGTAG TCGCACCAGC
5251      TGAAGCGATT CACAGATGTC TGCCTGTTCA TCCGCGTCCA GCTCGTTGAG
          ACTTCGCTAA GTGTCTACAG ACGGACAAGT AGGCGCAGGT CGAGCAACTC
5301      TTTCTCCAGA AGCGTTAATG TCTGGCTTCT GATAAAGCGG GCCATGTTAA
          AAAGAGGTCT TCGCAATTAC AGACCGAAGA CTATTTCGCC CGGTACAATT
5351      GGGCGGTTTT TTCCTGTTTG GTCACTGATG CCTCCGTGTA AGGGGGATTT
          CCCGCCAAAA AAGGACAAAC CAGTGACTAC GGAGGCACAT TCCCCCTAAA
5401      CTGTTCATGG GGGTAATGAT ACCGATGAAA CGAGAGAGGA TGCTCACGAT
          GACAAGTACC CCCATTACTA TGGCTACTTT GCTCTCTCCT ACGAGTGCTA
5451      ACGGGTTACT GATGATGAAC ATGCCCGGTT ACTGGAACGT TGTGAGGGTA
          TGCCCAATGA CTACTACTTG TACGGGCCAA TGACCTTGCA ACACTCCCAT
5501      AACAACTGGC GGTATGGATG CGGCGGGACC AGAGAAAAAT CACTCAGGGT
          TTGTTGACCG CCATACCTAC GCCGCCCTGG TCTCTTTTTA GTGAGTCCCA
5551      CAATGCCAGC GCTTCGTTAA TACAGATGTA GGTGTTCCAC AGGGTAGCCA
          GTTACGGTCG CGAAGCAATT ATGTCTACAT CCACAAGGTG TCCCATCGGT
5601      GCAGCATCCT GCGATGCAGA TCCGGAACAT AATGGTGCAG GGCGCTGACT
          CGTCGTAGGA CGCTACGTCT AGGCCTTGTA TTACCACGTC CCGCGACTGA
5651      TCCGCGTTTC CAGACTTTAC GAAACACGGA AACCGAAGAC CATTCATGTT
          AGGCGCAAAG GTCTGAAATG CTTTGTGCCT TTGGCTTCTG GTAAGTACAA
5701      GTTGCTCAGG TCGCAGACGT TTTGCAGCAG CAGTCGCTTC ACGTTCGCTC
          CAACGAGTCC AGCGTCTGCA AAACGTCGTC GTCAGCGAAG TGCAAGCGAG
5751      GCGTATCGGT GATTCATTCT GCTAACCAGT AAGGCAACCC CGCCAGCCTA
          CGCATAGCCA CTAAGTAAGA CGATTGGTCA TTCCGTTGGG GCGGTCGGAT
5801      GCCGGGTCCT CAACGACAGG AGCACGATCA TGCGCACCCG TGGCCAGGAC
          CGGCCCAGGA GTTGCTGTCC TCGTGCTAGT ACGCGTGGGC ACCGGTCCTG
                    AvaI
5851      CCAACGCTGC CCGAGATGCG CCGCGTGCGG CTGCTGGAGA TGGCGGACGC
          GGTTGCGACG GGCTCTACGC GGCGCACGCC GACGACCTCT ACCGCCTGCG
5901      GATGGATATG TTCTGCCAAG GGTTGGTTTG CGCATTCACA GTTCTCCGCA
          CTACCTATAC AAGACGGTTC CCAACCAAAC GCGTAAGTGT CAAGAGGCGT
                                          BsmI
```

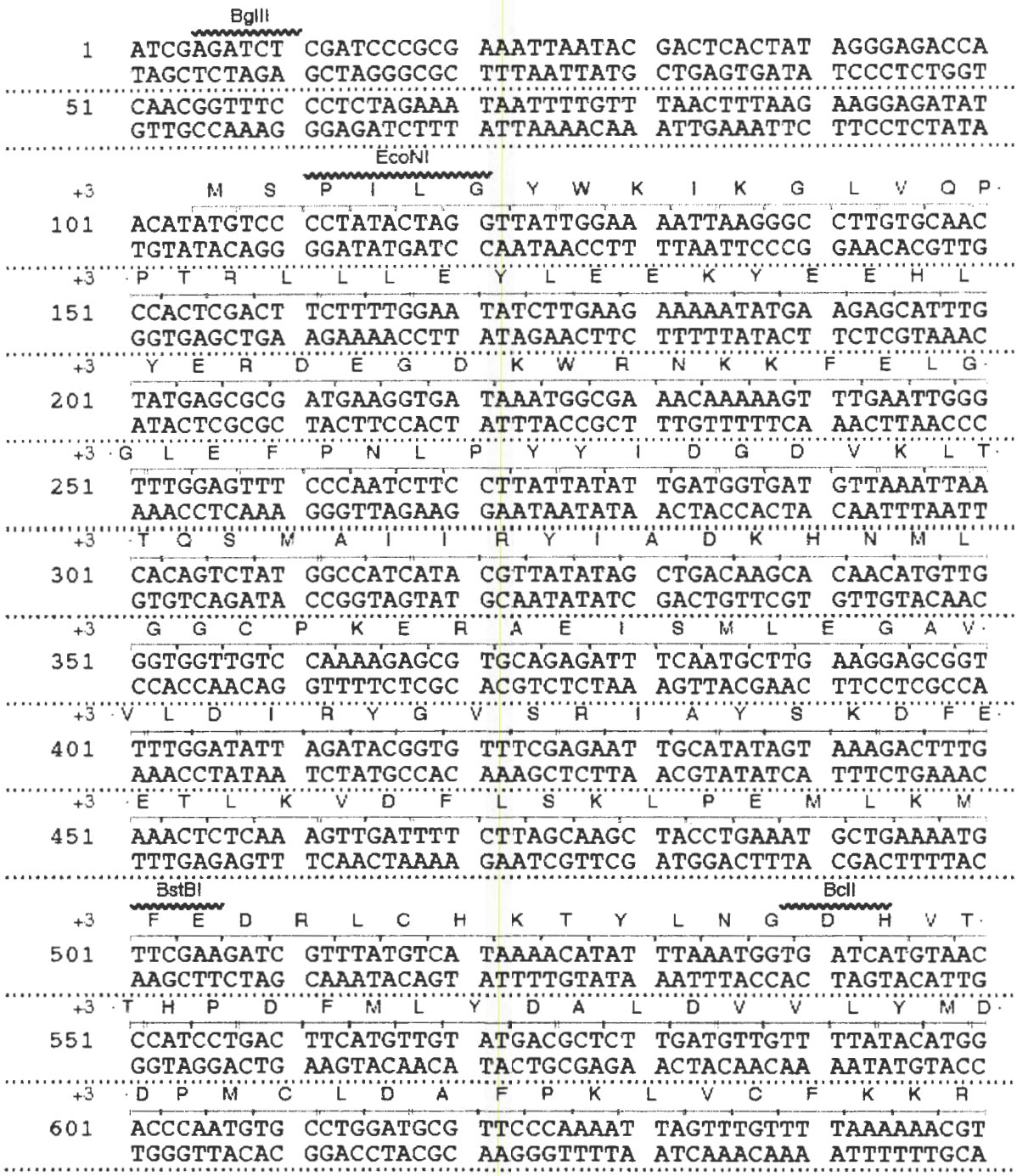
pDEST15+SSI

```
5951   AGAATTGATT GGCTCCAATT CTTGGAGTGG TGAATCCGTT AGCGAGGTGC
       TCTTAACTAA CCGAGGTTAA GAACCTCACC ACTTAGGCAA TCGCTCCACG
6001   CGCCGGCTTC CATTCAGGTC GAGGTGGCCC GGCTCCATGC ACCGCGACGC
       GCGGCCGAAG GTAAGTCCAG CTCCACCGGG CCGAGGTACG TGGCGCTGCG
6051   AACGCGGGGA GGCAGACAAG GTATAGGGCG GCGCCTACAA TCCATGCCAA
       TTGCGCCCCT CCGTCTGTTC CATATCCCGC CGCGGATGTT AGGTACGGTT
6101   CCCGTTCCAT GTGCTCGCCG AGGCGGCATA AATCGCCGTG ACGATCAGCG
       GGGCAAGGTA CACGAGCGGC TCCGCCGTAT TTAGCGGCAC TGCTAGTCGC
6151   GTCCAGTGAT CGAAGTTAGG CTGGTAAGAG CCGCGAGCGA TCCTTGAAGC
       CAGGTCACTA GCTTCAATCC GACCATTCTC GGCGCTCGCT AGGAACTTCG
6201   TGTCCCTGAT GGTCGTCATC TACCTGCCTG GACAGCATGG CCTGCAACGC
       ACAGGGACTA CCAGCAGTAG ATGGACGGAC CTGTCGTACC GGACGTTGCG
6251   GGGCATCCCG ATGCCGCCGG AAGCGAGAAG AATCATAATG GGGAAGGCCA
       CCCGTAGGGC TACGGCGGCC TTCGCTCTTC TTAGTATTAC CCCTTCCGGT
                 NruI
6301   TCCAGCCTCG CGTCGCGAAC GCCAGCAAGA CGTAGCCCAG CGCGTCGGCC
       AGGTCGGAGC GCAGCGCTTG CGGTCGTTCT GCATCGGGTC GCGCAGCCGG
6351   GCCATGCCGG CGATAATGGC CTGCTTCTCG CCGAAACGTT TGGTGGCGGG
       CGGTACGGCC GCTATTACCG GACGAAGAGC GGCTTTGCAA ACCACCGCCC
6401   ACCAGTGACG AAGGCTTGAG CGAGGGCGTG CAAGATTCCG AATACCGCAA
       TGGTCACTGC TTCCGAACTC GCTCCCGCAC GTTCTAAGGC TTATGGCGTT
6451   GCGACAGGCC GATCATCGTC GCGCTCCAGC GAAAGCGGTC CTCGCCGAAA
       CGCTGTCCGG CTAGTAGCAG CGCGAGGTCG CTTTCGCCAG GAGCGGCTTT
6501   ATGACCCAGA GCGCTGCCGG CACCTGTCCT ACGAGTTGCA TGATAAAGAA
       TACTGGGTCT CGCGACGGCC GTGGACAGGA TGCTCAACGT ACTATTTCTT
6551   GACAGTCATA AGTGCGGCGA CGATAGTCAT GCCCCGCGCC CACCGGAAGG
       CTGTCAGTAT TCACGCCGCT GCTATCAGTA CGGGGCGCGG GTGGCCTTCC
                                                  PvuI
6601   AGCTGACTGG GTTGAAGGCT CTCAAGGGCA TCGGTCGATC GACGCTCTCC
       TCGACTGACC CAACTTCCGA GAGTTCCCGT AGCCAGCTAG CTGCGAGAGG
                      EcoNI
6651   CTTATGCGAC TCCTGCATTA GGAAGCAGCC CAGTAGTAGG TTGAGGCCGT
       GAATACGCTG AGGACGTAAT CCTTCGTCGG GTCATCATCC AACTCCGGCA
6701   TGAGCACCGC CGCCGCAAGG AATGGTGCAT GCAAGGAGAT GGCGCCCAAC
       ACTCGTGGCG GCGGCGTTCC TTACCACGTA CGTTCCTCTA CCGCGGGTTG
                                                        BspHI
6751   AGTCCCCCGG CCACGGGGCC TGCCACCATA CCCACGCCGA AACAAGCGCT
       TCAGGGGGCC GGTGCCCCGG ACGGTGGTAT GGGTGCGGCT TTGTTCGCGA
       BspHI
6801   CATGAGCCCG AAGTGGCGAG CCCGATCTTC CCCATCGGTG ATGTCGGCGA
       GTACTCGGGC TTCACCGCTC GGGCTAGAAG GGGTAGCCAC TACAGCCGCT
```

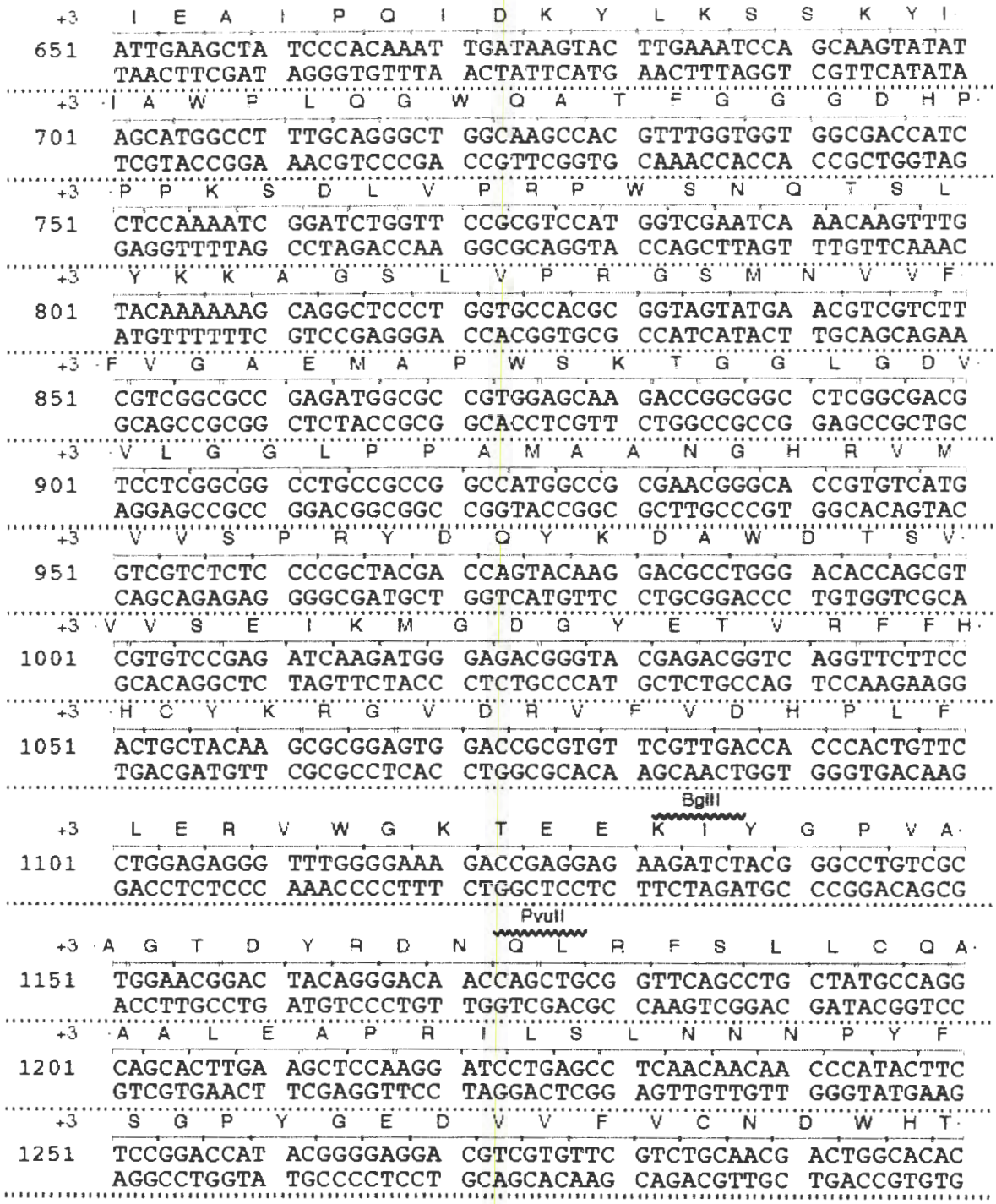
pDEST15+SSI

```
  6851   TATAGGCGCC AGCAACCGCA CCTGTGGCGC CGGTGATGCC GGCCACGATG
         ATATCCGCGG TCGTTGGCGT GGACACCGCG GCCACTACGG CCGGTGCTAC
  6901   CGTCCGGCGT AGAGG
         GCAGGCCGCA TCTCC
```

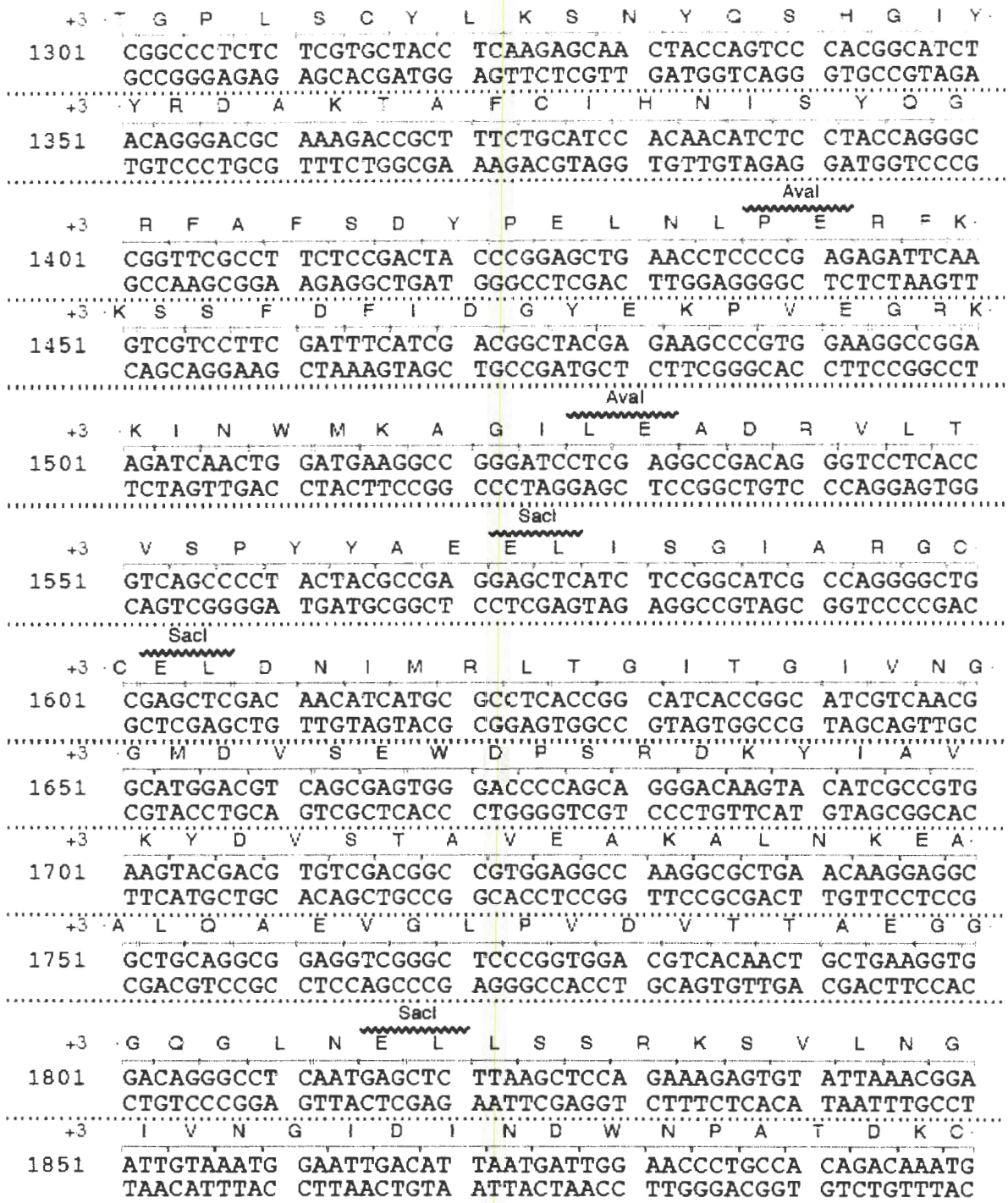
pDEST15+GBSSL+SSI #3



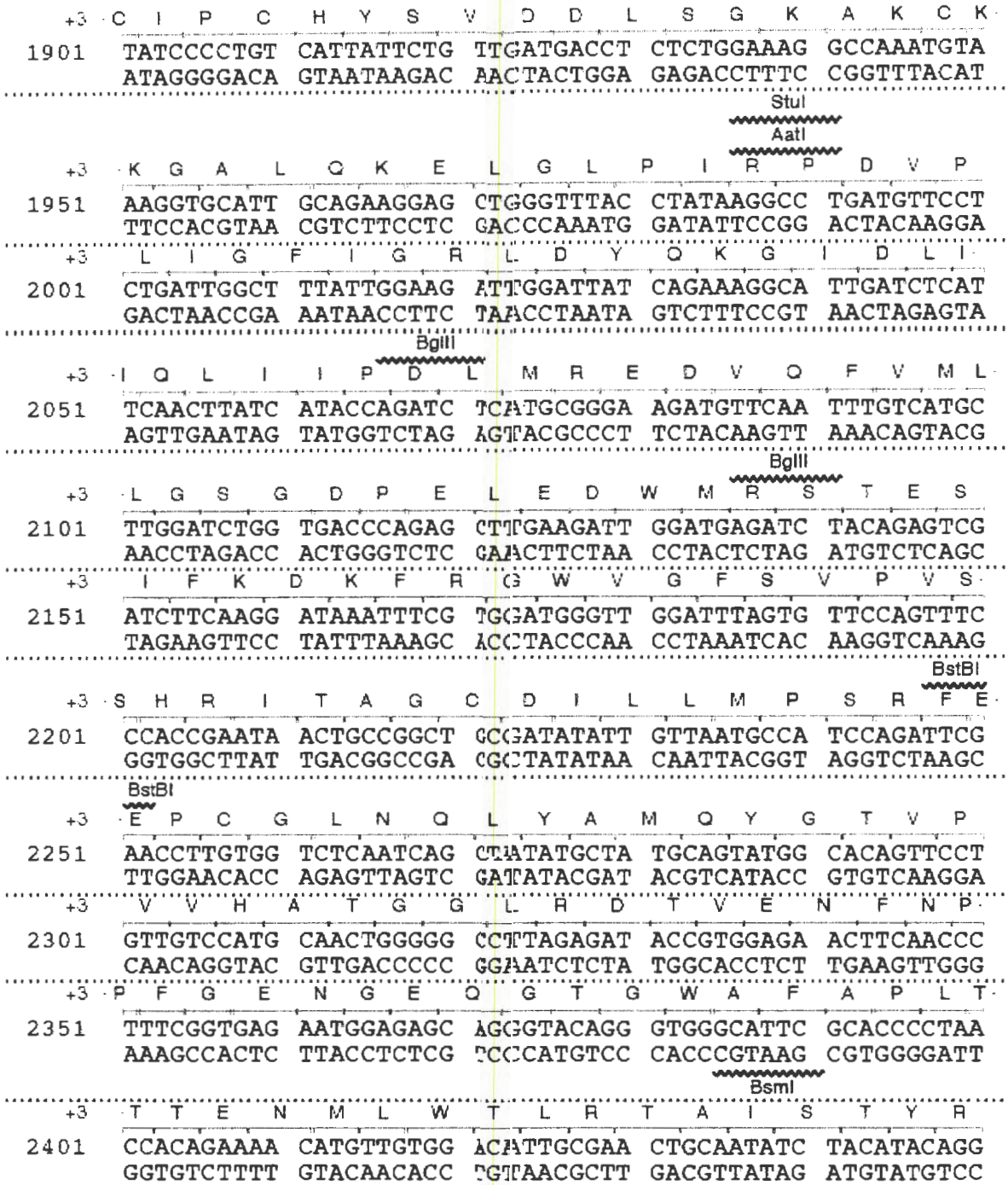
pDEST15+GBSSL+SSI #3



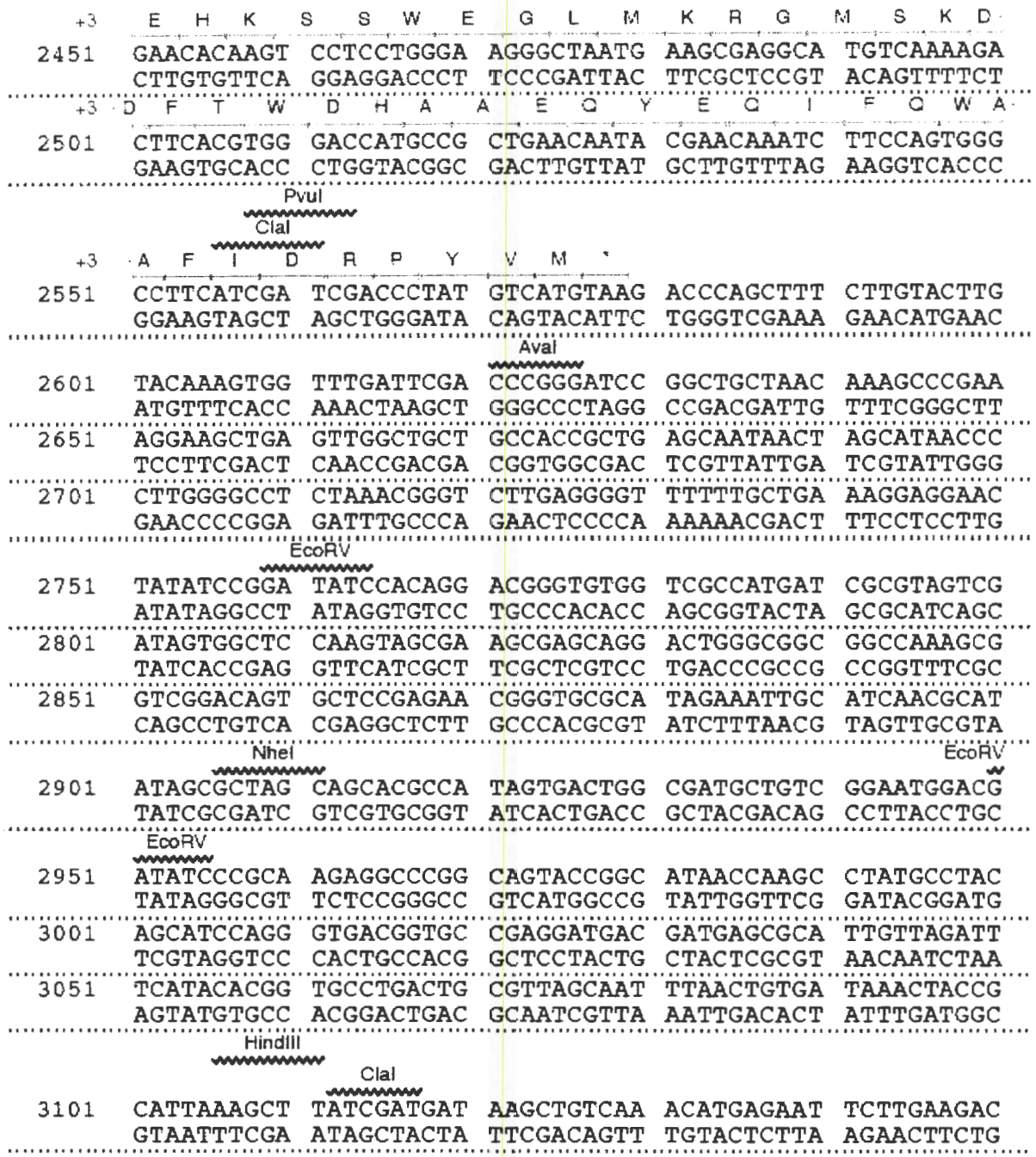
pDEST15+GBSSL+SSI #3



pDEST15+GBSSI+SSI #3



pDEST15+GBSSL+SSI #3



pDEST15+GESSL+SSI #3

```
                                                                  BspHI
3151  GAAAGGGCCT CGTGATACGC CTATTTTTAT AGGTTAATGT CATGATAATA
      CTTTCCCGGA GCACTATGCG GATAAAAATA TCCAATTACA GTACTATTAT
3201  ATGGTTTCTT AGACGTCAGG TGGCACTTTT CGGGGAAATG TGCGCGGAAC
      TACCAAAGAA TCTGCAGTCC ACCGTGAAAA GCCCCTTTAC ACGCGCCTTG
                                                        BspHI
3251  CCCTATTTGT TTATTTTTCT AAATACATTC AAATATGTAT CCGCTCATGA
      GGGATAAACA AATAAAAAGA TTTATGTAAG TTTATACATA GGCGAGTACT
3301  GACAATAACC CTGATAAATG CTTCAATAAT ATTGAAAAAG GAAGAGTATG
      CTGTTATTGG GACTATTTAC GAAGTTATTA TAACTTTTTC CTTCTCATAC
3351  AGTATTCAAC ATTTCCGTGT CGCCCTTATT CCCTTTTTTG CGGCATTTTG
      TCATAAGTTG TAAAGGCACA GCGGGAATAA GGGAAAAAAC GCCGTAAAAC
3401  CCTTCCTGTT TTTGCTCACC CAGAAACGCT GGTGAAAGTA AAAGATGCTG
      GGAAGGACAA AAACGAGTGG GTCTTTGCGA CCACTTTCAT TTTCTACGAC
                  ApaLI
3451  AAGATCAGTT GGGTGCACGA GTGGGTTACA TCGAACTGGA TCTCAACAGC
      TTCTAGTCAA CCCACGTGCT CACCCAATGT AGCTTGACCT AGAGTTGTCG
3501  GGTAAGATCC TTGAGAGTTT TCGCCCCGAA GAACGTTTTC CAATGATGAG
      CCATTCTAGG AACTCTCAAA AGCGGGGCTT CTTGCAAAAG GTTACTACTC
3551  CACTTTTAAA GTTCTGCTAT GTGGCGCGGT ATTATCCCGT GTTGACGCCG
      GTGAAAATTT CAAGACGATA CACCGCGCCA TAATAGGGCA CAACTGCGGC
3601  GGCAAGAGCA ACTCGGTCGC CGCATACACT ATTCTCAGAA TGACTTGGTT
      CCGTTCTCGT TGAGCCAGCG GCGTATGTGA TAAGAGTCTT ACTGAACCAA
3651  GAGTACTCAC CAGTCACAGA AAAGCATCTT ACGGATGGCA TGACAGTAAG
      CTCATGAGTG GTCAGTGTCT TTTCGTAGAA TGCCTACCGT ACTGTCATTC
3701  AGAATTATGC AGTGCTGCCA TAACCATGAG TGATAACACT GCGGCCAACT
      TCTTAATACG TCACGACGGT ATTGGTACTC ACTATTGTGA CGCCGGTTGA
                  PvuI
3751  TACTTCTGAC AACGATCGGA GGACCGAAGG AGCTAACCGC TTTTTTGCAC
      ATGAAGACTG TTGCTAGCCT CCTGGCTTCC TCGATTGGCG AAAAAACGTG
3801  AACATGGGGG ATCATGTAAC TCGCCTTGAT CGTTGGGAAC CGGAGCTGAA
      TTGTACCCCC TAGTACATTG AGCGGAACTA GCAACCCTTG GCCTCGACTT
3851  TGAAGCCATA CCAAACGACG AGCGTGACAC CACGATGCCT GCAGCAATGG
      ACTTCGGTAT GGTTTGCTGC TCGCACTGTG GTGCTACGGA CGTCGTTACC
3901  CAACAACGTT GCGCAAACTA TTAACTGGCG AACTACTTAC TCTAGCTTCC
      GTTGTTGCAA CGCGTTTGAT AATTGACCGC TTGATGAATG AGATCGAAGG
3951  CGGCAACAAT TAATAGACTG GAATGGAGGCG GATAAAGTTG CAGGACCACT
      GCCGTTGTTA ATTATCTGAC CTTACCTCCGC CTATTTCAAC GTCCTGGTGA
4001  TCTGCGCTCG GCCCTTCCGG CTGGCTGGTT TATTGCTGAT AAATCTGGAG
      AGACGCGAGC CGGGAAGGCC GACCGACCAA ATAACGACTA TTTAGACCTC
4051  CCGGTGAGCG TGGGTCTCGC GGTATCATTG CAGCACTGGG GCCAGATGGT
      GGCCACTCGC ACCCAGAGCG CCATAGTAAC GTCGTGACCC CGGTCTACCA
```

pDEST15+GBSSL+SSI #3

```
4101  AAGCCCTCCC GTATCGTAGT TATCTACACG ACGGGGAGTC AGGCAACTAT
      TTCGGGAGGG CATAGCATCA ATAGATGTGC TGCCCCTCAG TCCGTTGATA
4151  GGATGAACGA AATAGACAGA TCGCTGAGAT AGGTGCCTCA CTGATTAAGC
      CCTACTTGCT TTATCTGTCT AGCGACTCTA TCCACGGAGT GACTAATTCG
4201  ATTGGTAACT GTCAGACCAA GTTTACTCAT ATATACTTTA GATTGATTTA
      TAACCATTGA CAGTCTGGTT CAAATGAGTA TATATGAAAT CTAACTAAAT
4251  AAACTTCATT TTTAATTTAA AAGGATCTAG GTGAAGATCC TTTTTGATAA
      TTTGAAGTAA AAATTAAATT TTCCTAGATC CACTTCTAGG AAAAACTATT
            BspHI
4301  TCTCATGACC AAAATCCCTT AACGTGAGTT TTCGTTCCAC TGAGCGTCAG
      AGAGTACTGG TTTTAGGGAA TTGCACTCAA AAGCAAGGTG ACTCGCAGTC
4351  ACCCCGTAGA AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC
      TGGGGCATCT TTTCTAGTTT CCTAGAAGAA CTCTAGGAAA AAAAGACGCG
4401  GTAATCTGCT GCTTGCAAAC AAAAAAACCA CCGCTACCAG CGGTGGTTTG
      CATTAGACGA CGAACGTTTG TTTTTTTGGT GGCGATGGTC GCCACCAAAC
4451  TTTGCCGGAT CAAGAGCTAC CAACTCTTTT TCCGAAGGTA ACTGGCTTCA
      AAACGGCCTA GTTCTCGATG GTTGAGAAAA AGGCTTCCAT TGACCGAAGT
4501  GCAGAGCGCA GATACCAAAT ACTGTCCTTC TAGTGTAGCC GTAGTTAGGC
      CGTCTCGCGT CTATGGTTTA TGACAGGAAG ATCACATCGG CATCAATCCG
4551  CACCACTTCA AGAACTCTGT AGCACCGCCT ACATACCTCG CTCTGCTAAT
      GTGGTGAAGT TCTTGAGACA TCGTGGCGGA TGTATGGAGC GAGACGATTA
4601  CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT CTTACCGGGT
      GGACAATGGT CACCGACGAC GGTCACCGCT ATTCAGCACA GAATGGCCCA
4651  TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG
      ACCTGAGTTC TGCTATCAAT GGCCTATTCC GCGTCGCCAG CCCGACTTGC
                  ApaLI
4701  GGGGGTTCGT GCACACAGCC CAGCTTGGAG CGAACGACCT ACACCGAACT
      CCCCCAAGCA CGTGTGTCGG GTCGAACCTC GCTTGCTGGA TGTGGCTTGA
4751  GAGATACCTA CAGCGTGAGC TATGAGAAAG CGCCACGCTT CCCGAAGGGA
      CTCTATGGAT GTCGCACTCG ATACTCTTTC GCGGTGCGAA GGGCTTCCCT
4801  GAAAGGCGGA CAGGTATCCG GTAAGCGGCA GGGTCGGAAC AGGAGAGCGC
      CTTTCCGCCT GTCCATAGGC CATTCGCCGT CCCAGCCTTG TCCTCTCGCG
4851  ACGAGGGAGC TTCCAGGGGG AAACGCCTGG TATCTTTATA GTCCTGTCGG
      TGCTCCCTCG AAGGTCCCCC TTTGCGGACC ATAGAAATAT CAGGACAGCC
4901  GTTTCGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC TCGTCAGGGG
      CAAAGCGGTG GAGACTGAAC TCGCAGCTAA AAACACTACG AGCAGTCCCC
4951  GGCGGAGCCT ATGGAAAAAC GCCAGCAACG CGGCCTTTTT ACGGTTCCTG
      CCGCCTCGGA TACCTTTTTG CGGTCGTTGC GCCGGAAAAA TGCCAAGGAC
5001  GCCTTTTGCT GGCCTTTTGC TCACATGTTC TTTCCTGCGT TATCCCCTGA
      CGGAAAACGA CCGGAAAACG AGTGTACAAG AAAGGACGCA ATAGGGGACT
5051  TTCTGTGGAT AACCGTATTA CCGCCTTTGA GTGAGCTGAT ACCGCTCGCC
      AAGACACCTA TTGGCATAAT GGCGGAAACT CACTCGACTA TGGCGAGCGG
```

pDEST15+GBSSL+SSI #3

```
5101   GCAGCCGAAC GACCGAGCGC AGCGAGTCAG TGAGCGAGGA AGCGGAAGAG
       CGTCGGCTTG CTGGCTCGCG TCGCTCAGTC ACTCGCTCCT TCGCCTTCTC
5151   CGCCTGATGC GGTATTTTCT CCTTACGCAT CTGTGCGGTA TTTCACACCG
       GCGGACTACG CCATAAAAGA GGAATGCGTA GACACGCCAT AAAGTGTGGC
                  ApaLI
5201   CATATATGGT GCACTCTCAG TACAATCTGC TCTGATGCCG CATAGTTAAG
       GTATATACCA CGTGAGAGTC ATGTTAGACG AGACTACGGC GTATCAATTC
5251   CCAGTATACA CTCCGCTATC GCTACGTGAC TGGGTCATGG CTGCGCCCCG
       GGTCATATGT GAGGCGATAG CGATGCACTG ACCCAGTACC GACGCGGGGC
5301   ACACCCGCCA ACACCCGCTG ACGCGCCCTG ACGGGCTTGT CTGCTCCCGG
       TGTGGGCGGT TGTGGGCGAC TGCGCGGGAC TGCCCGAACA GACGAGGGCC
5351   CATCCGCTTA CAGACAAGCT GTGACCGTCT CCGGGAGCTG CATGTGTCAG
       GTAGGCGAAT GTCTGTTCGA CACTGGCAGA GGCCCTCGAC GTACACAGTC
                                                  PvuII
5401   AGGTTTTCAC CGTCATCACC GAAACGCGCG AGGCAGCTGC GGTAAAGCTC
       TCCAAAAGTG GCAGTAGTGG CTTTGCGCGC TCCGTCGACG CCATTTCGAG
5451   ATCAGCGTGG TCGTGAAGCG ATTCACAGAT GTCTGCCTGT TCATCCGCGT
       TAGTCGCACC AGCACTTCGC TAAGTGTCTA CAGACGGACA AGTAGGCGCA
5501   CCAGCTCGTT GAGTTTCTCC AGAAGCGTTA ATGTCTGGCT TCTGATAAAG
       GGTCGAGCAA CTCAAAGAGG TCTTCGCAAT TACAGACCGA AGACTATTTC
5551   CGGGCCATGT TAAGGGCGGT TTTTTCCTGT TTGGTCACTG ATGCCTCCGT
       GCCCGGTACA ATTCCCGCCA AAAAAGGACA AACCAGTGAC TACGGAGGCA
5601   GTAAGGGGGA TTTCTGTTCA TGGGGGTAAT GATACCGATG AAACGAGAGA
       CATTCCCCCT AAAGACAAGT ACCCCCATTA CTATGGCTAC TTTGCTCTCT
5651   GGATGCTCAC GATACGGGTT ACTGATGATG AACATGCCCG GTTACTGGAA
       CCTACGAGTG CTATGCCCAA TGACTACTAC TTGTACGGGC CAATGACCTT
5701   CGTTGTGAGG GTAAACAACT GGCGGTATGG ATGCGGCGGG ACCAGAGAAA
       GCAACACTCC CATTTGTTGA CCGCCATACC TACGCCGCCC TGGTCTCTTT
5751   AATCACTCAG GGTCAATGCC AGCGCTTCGT TAATACAGAT GTAGGTGTTC
       TTAGTGAGTC CCAGTTACGG TCGCGAAGCA ATTATGTCTA CATCCACAAG
5801   CACAGGGTAG CCAGCAGCAT CCTGCGATGC AGATCCGGAA CATAATGGTG
       GTGTCCCATC GGTCGTCGTA GGACGCTACG TCTAGGCCTT GTATTACCAC
5851   CAGGGCGCTG ACTTCCGCGT TTCCAGACTT TACGAAACAC GGAAACCGAA
       GTCCCGCGAC TGAAGGCGCA AAGGTCTGAA ATGCTTTGTG CCTTTGGCTT
5901   GACCATTCAT GTTGTTGCTC AGGTCGCAGA CGTTTTGCAG CAGCAGTCGC
       CTGGTAAGTA CAACAACGAG TCCAGCGTCT GCAAAACGTC GTCGTCAGCG
5951   TTCACGTTCG CTCGCGTATC GGTGATTCAT TCTGCTAACC AGTAAGGCAA
       AAGTGCAAGC GAGCGCATAG CCACTAAGTA AGACGATTGG TCATTCCGTT
6001   CCCCGCCAGC CTAGCCGGGT CCTCAACGAC AGGAGCACGA TCATGCGCAC
       GGGGCGGTCG GATCGGCCCA GGAGTTGCTG TCCTCGTGCT AGTACGCGTG
```

pDEST15+GBSSL+SSI #3

```
                                        AvaI
6051  CCGTGGCCAG GACCCAACGC TGCCCGAGAT GCGCCGCGTG CGGCTGCTGG
      GGCACCGGTC CTGGGTTGCG ACGGGCTCTA CGCGGCGCAC GCCGACGACC
6101  AGATGGCGGA CGCGATGGAT ATGTTCTGCC AAGGGTTGGT TTGCGCATTC
      TCTACCGCCT GCGCTACCTA TACAAGACGG TTCCCAACCA AACGCGTAAG
                                                        BsmI
6151  ACAGTTCTCC GCAAGAATTG ATTGGCTCCA ATTCTTGGAG TGGTGAATCC
      TGTCAAGAGG CGTTCTTAAC TAACCGAGGT TAAGAACCTC ACCACTTAGG
6201  GTTAGCGAGG TGCCGCCGGC TTCCATTCAG GTCGAGGTGG CCCGGCTCCA
      CAATCGCTCC ACGGCGGCCG AAGGTAAGTC CAGCTCCACC GGGCCGAGGT
6251  TGCACCGCGA CGCAACGCGG GGAGGCAGAC AAGGTATAGG GCGGCGCCTA
      ACGTGGCGCT GCGTTGCGCC CCTCCGTCTG TTCCATATCC CGCCGCGGAT
6301  CAATCCATGC CAACCCGTTC CATGTGCTCG CCGAGGCGGC ATAAATCGCC
      GTTAGGTACG GTTGGGCAAG GTACACGAGC GGCTCCGCCG TATTTAGCGG
6351  GTGACGATCA GCGGTCCAGT GATCGAAGTT AGGCTGGTAA GAGCCGCGAG
      CACTGCTAGT CGCCAGGTCA CTAGCTTCAA TCCGACCATT CTCGGCGCTC
6401  CGATCCTTGA AGCTGTCCCT GATGGTCGTC ATCTACCTGC CTGGACAGCA
      GCTAGGAACT TCGACAGGGA CTACCAGCAG TAGATGGACG GACCTGTCGT
6451  TGGCCTGCAA CGCGGGCATC CCGATGCCGC CGGAAGCGAG AAGAATCATA
      ACCGGACGTT GCGCCCGTAG GGCTACGGCG GCCTTCGCTC TTCTTAGTAT
                                        NruI
6501  ATGGGGAAGG CCATCCAGCC TCGCGTCGCG AACGCCAGCA AGACGTAGCC
      TACCCCTTCC GGTAGGTCGG AGCGCAGCGC TTGCGGTCGT TCTGCATCGG
6551  CAGCGCGTCG GCCGCCATGC CGGCGATAAT GGCCTGCTTC TCGCCGAAAC
      GTCGCGCAGC CGGCGGTACG GCCGCTATTA CCGGACGAAG AGCGGCTTTG
6601  GTTTGGTGGC GGGACCAGTG ACGAAGGCTT GAGCGAGGGC GTGCAAGATT
      CAAACCACCG CCCTGGTCAC TGCTTCCGAA CTCGCTCCCG CACGTTCTAA
6651  CCGAATACCG CAAGCGACAG GCCGATCATC GTCGCGCTCC AGCGAAAGCG
      GGCTTATGGC GTTCGCTGTC CGGCTAGTAG CAGCGCGAGG TCGCTTTCGC
6701  GTCCTCGCCG AAAATGACCC AGAGCGCTGC CGGCACCTGT CCTACGAGTT
      CAGGAGCGGC TTTTACTGGG TCTCGCGACG GCCGTGGACA GGATGCTCAA
6751  GCATGATAAA GAAGACAGTC ATAAGTGCGG CGACGATAGT CATGCCCCGC
      CGTACTATTT CTTCTGTCAG TATTCACGCC GCTGCTATCA GTACGGGGCG
                                                        PvuI
6801  GCCCACCGGA AGGAGCTGAC TGGGTTGAAG GCTCTCAAGG GCATCGGTCG
      CGGGTGGCCT TCCTCGACTG ACCCAACTTC CGAGAGTTCC CGTAGCCAGC
      PvuI                              EcoNI
6851  ATCGACGCTC TCCCTTATGC GACTCCTGCA TTAGGAAGCA GCCCAGTAGT
      TAGCTGCGAG AGGGAATACG CTGAGGACGT AATCCTTCGT CGGGTCATCA
6901  AGGTTGAGGC CGTTGAGCAC CGCCGCCGCA AGGAATGGTG CATGCAAGGA
      TCCAACTCCG GCAACTCGTG GCGGCGGCGT TCCTTACCAC GTACGTTCCT
```

pDEST15+GBSSL+SSI #3

```
6951   GATGGCGCCC AACAGTCCCC CGGCCACGGG GCCTGCCACC ATACCCACGC
       CTACCGCGGG TTGTCAGGGG GCCGGTGCCC CGGACGGTGG TATGGGTGCG
                          BspHI
7001   CGAAACAAGC GCTCATGAGC CCGAAGTGGC GAGCCCGATC TTCCCCATCG
       GCTTTGTTCG CGAGTACTCG GGCTTCACCG CTCGGGCTAG AAGGGGTAGC
7051   GTGATGTCGG CGATATAGGC GCCAGCAACC GCACCTGTGG CGCCGGTGAT
       CACTACAGCC GCTATATCCG CGGTCGTTGG CGTGGACACC GCGGCCACTA
7101   GCCGGCCACG ATGCGTCCGG CGTAGAGG
       CGGCCGGTGC TACGCAGGCC GCATCTCC
```

APPENDIX B. ANALYSIS OF GRANULE BOUND PROTEINS IN AN ALLELIC SERIES OF *SUGARY2* MUTANT STARCHES

Introduction

In higher plants starch is formed for the storage of a large amount of glucose, or reduced carbon, formed in photosynthesis. An array of enzymes are involved in the starch synthesis including adenosine 5' diphosphate glucose pyrophosphorylase (AGP), starch synthases (SS's), branching enzymes (BE's), debranching enzymes (DBE's), starch phosphorylase (SP), and disproportionating enzyme (D-enyme). AGP converts glucose-1-phosophate to adenosine 5' diphosphate glucose (ADPG) and is the first committed step in the synthesis of starch. SS's take the ADPG and add it on to the end of a pre-existing glucan via α-1, 4 linkage. BE's introduce branch points to a glucan through the cleavage of a region of α-1, 4 linked glucan and transfers it to an acceptor chain forming an α-1, 6 linkage. DBE's hydrolyze α-1, 6 glucan linkages and are thought to assist in the organization of glucan chains in order to create a semi-crystalline granule. The exact roles of SP and D-enzyme are unresolved, but current research suggests that they are also involved the starch synthesis mechanism.

The focus of this study is on the enzymes that become entrapped or are associated with the granule fraction of starch. In wild type maize, there are three enzymes physically associated with the granule including granule bound starch synthase (GBSS), more than 85% of starch synthase I (SSI), and more than 45% of branching enzyme II (BEII) [1]. GBSS and SSI are both SS's and are involved in the formation of amylose and amylopectin respectively. There are three other SS's in maize that function in the creation of amylopectin including starch synthase IIa (SSIIa), starch synthase IIb (SSIIb), and starch synthase III (SSIII). BEII, later found to be BEIIa and BEIIb, and branching enzyme I (BEI) are the only known BE's in maize starch and involved in the formation of branch points in amylopectin.

Lost of one or more of these enzymes can have a significant effect on granule morphology. In maize *sugary2* mutants lack the activity of SSIIa and causes the amylopectin to have an increased abundance of short glucan chains and a proportional decrease in intermediate length chains [2]. This altered granule structure is similar to the effects of SSII deficiency in other species [3-8]. Morell *et al.* [3] found in barley *sex6* mutants, lacking SSIIa activity, that not only chain lengths were altered but the enzymes associated with the granule in wild type were lessened or no longer found associated with the granule in the mutant including SSI, BEIIa, and BEIIb.

This report investigates the enzymes associated in maize starch granules of wild type (W64A) and five *su2* mutant plants. Each *su2* mutant varies in the level of severity of the mutation including *su2*-Ref, *su2*-1982, *su2*-19791, *su2*-5178, and *su2*-2279 in order from least severe to most severe [2]. The hypothesis is that the granule association of maize SSI, BEI, and BEIIa in the *su2* mutants will be lessened or no longer found as had been discovered in barley. Results indicate that there are no significant changes in the levels of BEI and BEIIa compared to wild type. Slight differences in SSI levels were noticed in the most severe *su2* mutants compared to wild type. These results indicate that maize SSIIa mutants do not show the same effects on enzyme association with the granule as barley.

Materials and Methods

Materials

Plants were field-grown in the summer at Iowa State University, or were grown in the greenhouse under standard light conditions with supplemental lighting on a 14/10-h day/night cycle and the homozygous *su2*-5178, *su2*-2279, and *su2*-Ref mutants were developed into the W64A inbred background as explained in Zhang *et al.* [2]. Two other *su2* mutants, *su2*-1982 and *su2*-19791, were obtained from the Maize Genetic Stock Center (Urbana, IL) whose alleles have not yet been back crossed into a standard background. Antibodies for maize SSI were a generous gift from Dr. Guan and Dr. Keeling (BASF-Plant

Sciences, Ames, IA). BEI and BEIIa antibodies were developed in our laboratory by Rebekah Marsh. Thermolysin (protease type X from *Bacillus thermoproteolyticus*: EC 3.4.24.4) was obtained from Sigma.

Starch granule isolation

Granule isolation methods were developed from the methods described by Mu-Forster *et al.* [1]. Starch granules were isolated from 5 g of 20 DAP maize endosperm by homogenizing in 10 ml of cold starch extraction buffer (SEB; 1.25 mM DTT, 10 mM EDTA, 10% glycerol, 50 mM Tris-HCl (pH 7.0)) in an ice cold mortar and pestle. Homogenate was filtered through two layers of Miracloth (pre-wet with a little SEB) into a centrifuge tube on ice. Filtrate was centrifuged at 15,000*g* for 15 min at 4°C. The delicate yellow gel-like layer on top of the white layer of sedimented starch was carefully removed with a spatula and discarded in order to obtain clean white granules. Then the pellet was washed twice with 5 ml cold SEB, twice with 5 ml cold 70% ethanol to remove low-molecular weight storage proteins, and twice with 10 ml of cold acetone. Finally, the starch pellet was dried with a gentle stream of air. At this point the starch is considered "crude" because it still has proteins associated with the outside of the granule.

Protease digestion and detergent extraction of starch granules

Methods involved in removing strongly adherent proteins associated with the surface of the granule are developed from the methods described by Mu-Forster *et al.* [1]. 50 mg of "crude" extracted dry starch granules were incubated with 100 µg of thermolysin and 5 mM $CaCl_2$ at 37°C for 30 min. The reaction was terminated by the addition of cold EDTA to 20 mM and centrifuged at 13,000*g* for 5 min at 4°C. Supernatant was removed and discarded leaving only the white starch pellet. Pellet was washed in the following order: 5 times with 1 ml cold water, once with 1 ml cold 2% SDS, 3 times with 1 ml cold water, once with 1 ml cold 70% ethanol, and once with 1 ml cold acetone. After all of the washing, pellet was dried under a gentle stream of air.

Protein extraction from starch granules

Again, the method to extract the proteins from the granule were developed from the methods described by Mu-Forster *et al.* [1]. 1 ml of SDS-PAGE sample buffer (3% SDS, 5% ß-mercaptoethanol, 10% glycerol, and 62.5 mM Tris-HCl (pH 6.9)) was added to 50 mg of starch granules and boiled for 15 min. Samples were allowed to cool and then were centrifuged for 5 min at 13,000*g* at room temperature. The supernatant containing extracted proteins was analyzed by SDS-PAGE.

Protein verification

Proteins were examined using standard biochemistry protocols for SDS-PAGE (10% acrylamide) by Coomassie stain and verified using immunodetection with antibodies for SSI, BEI, BEIIa, and GBSS [9]. Because the granule proteins could not be extracted without boiling in SDS loading buffer the protein concentration in solution could not be determined. However, exactly 50 mg of granules were used in each sample so the same amount relative amount of protein should be present depending on the amount of proteins that are starch granule bound.

Results & Discussion

In this report the maize granule bound proteins of wild type and *su2* mutants were investigated to detect any differences in the amount of enzyme entrapped in the granule. Lost of SSIIa in maize and other species has caused an alteration in starch granule morphology however differences in enzyme location had not been examined except in barley *sex6* mutants [3]. Morell et al. [3] showed a loss and/or decrease in BEIIa, BEIIb, and SSI in SSIIa mutants (see Figure 3). The hypothesis is that the granule association of maize SSI, BEI, and BEIIa in the *su2* mutants will be lessened or no longer found as had been discovered in barley.

Crude and clean starch granule proteins from wild type and *su2* mutant maize were extracted and examined using SDS-PAGE. Thermolysin treated granule proteins were compared with crude starch to visualize the loss of proteins associated with the granule surface. Successful treatment was indicated as a number of protein bands disappeared from thermolysin treated granules compared to the crude granules (see Figure 1).

The thermolysin treated granules were then used to determine variance between wild type and *su2* mutant proteins entrapped in the granule. Treatment with thermolysin is very important because it ensures that the only proteins being identified are those that are actually entrapped in the granule instead of including those that are simply associated with the granule surface. The Coomassie stain results of the wild type and five *su2* mutants were not as expected. The hypothesis was that there would be a loss of some proteins in the mutants compared to the wild type, but instead it appears as if the mutants have increase of intensity of protein bands most notably in the lower molecular weight molecules (see Figure 2a). It also appears as if there is another protein band around 60 kD in the *su2* mutants (see Figure 1 and 2a).

Immunodetection was used to further investigate protein levels of SSI, BEI, BEIIa, and GBSS in the granule. BEI, BEIIa, and GBSS did not show significant changes in intensity between the wild type and mutants (see Figure 2b and 2c). SSI showed a slight decrease only in the most severe mutant, *su2*-2279, compared to the wild type (see Figure 2d). GBSS is ruled out for the identity of the extra band in the *su2* mutants because the level of GBSS protein bound is the same amount in all of the starch types. Future research is needed to investigate the identity of this protein perhaps by N-terminal amino acid sequencing.

Previous research indicates that SSIIa does not associate with the granule in maize [1]. In barley SSII is associated with the granule and the soluble fractions of starch [10]. This may indicate a slight difference in either function or affinity of SSIIa in maize and SSIIa in barley. Conversely, granule morphology of SSIIa mutants in maize and barley show the same types of alterations in amylopectin [2,3]. Further research is required to investigate the

similarities and differences of an SSIIa mutation in multiple species in order to fully understand how SSIIa affects the location of other proteins in the granule fraction of starch.

References

1. Mu-Forster, C., et al., *Physical association of starch biosynthetic enzymes with starch granules of maize endosperm. Granule-associated forms of starch synthase I and starch branching enzyme II.* Plant Physiol, 1996. **111**(3): p. 821-9.
2. Zhang, X., et al., *Molecular characterization demonstrates that the Zea mays gene sugary2 codes for the starch synthase isoform SSIIa.* Plant Mol Biol, 2004. **54**(6): p. 865-79.
3. Morell, M.K., et al., *Barley sex6 mutants lack starch synthase IIa activity and contain a starch with novel properties.* Plant J, 2003. **34**(2): p. 173-85.
4. Yamamori, M., et al., *Genetic elimination of a starch granule protein, SGP-1, of wheat generates an altered starch with apparent high amylose.* Theor Appl Genet, 2000. **101**: p. 21-29.
5. Umemoto, T., et al., *Mapping of a gene responsible for the difference in amylopectin structure between japonica-type and indica-type rice varieties.* Theor Appl Genet, 2002. **104**(1): p. 1-8.
6. Craig, J., et al., *Mutations in the gene encoding starch synthase II profoundly alter amylopectin structure in pea embryos.* Plant Cell, 1998. **10**(3): p. 413-26.
7. Fontaine, T., et al., *Toward an understanding of the biogenesis of the starch granule. Evidence that Chlamydomonas soluble starch synthase II controls the synthesis of intermediate size glucans of amylopectin.* J Biol Chem, 1993. **268**(22): p. 16223-30.
8. Edwards, A., et al., *A combined reduction in activity of starch synthases II and III of potato has novel effects on the starch of tubers.* The Plant Journal, 1999. **17**(3): p. 251-261.

9. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.

10. Li, Z., et al., *The structural organisation of the gene encoding class II starch synthase of wheat and barley and the evolution of the genes encoding starch synthases in plants.* Funct Integr Genomics, 2003. **3**(1-2): p. 76-85.

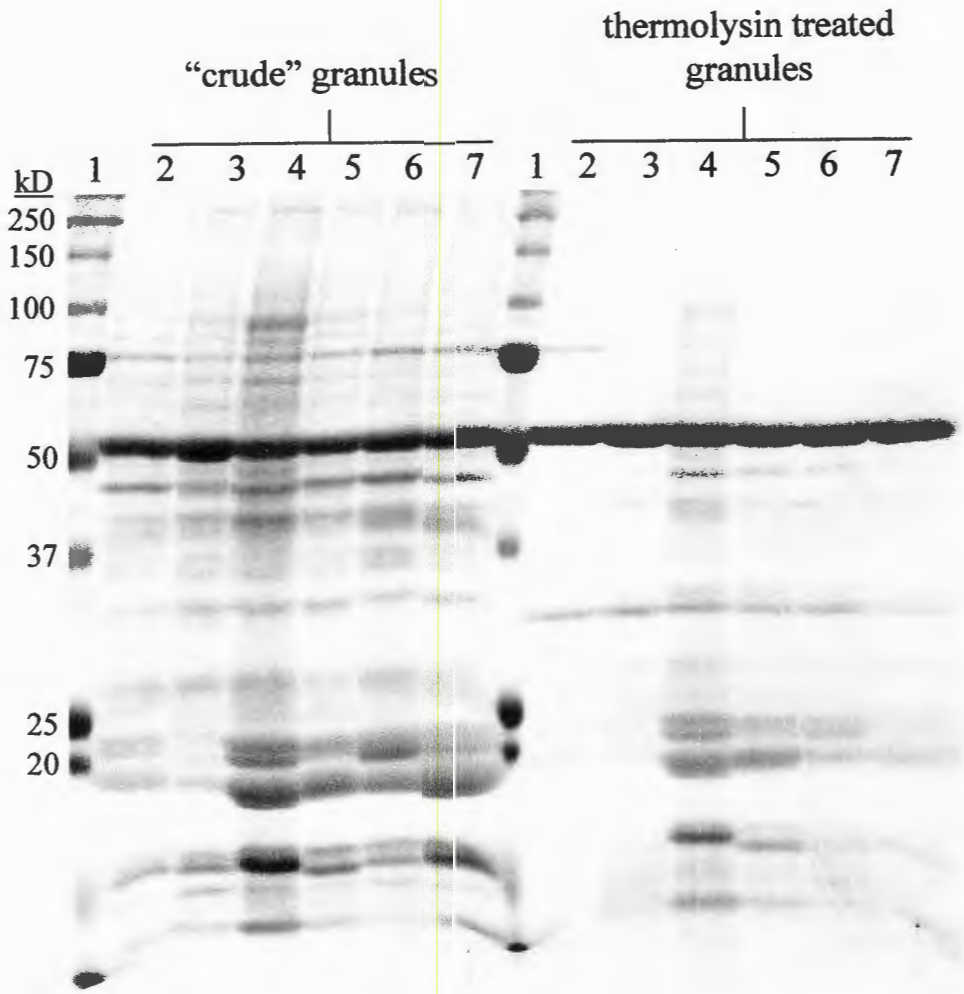


Figure 1: Coomassie stain of "crude" and thermolysin treated granules of wild type and *su2* mutant starch granules. Notice the loss or lessening of some of the protein bands going from the "crude" starch to the thermolysin treated starch in all of the samples. Lanes are as follows: 1-molecular weight marker, 2-wild type (W64A), 3-*su2*-2279, 4-*su2*-5178, 5-*su2*-1982, 6-*su2*-19791, and 7-*su2*-Ref.

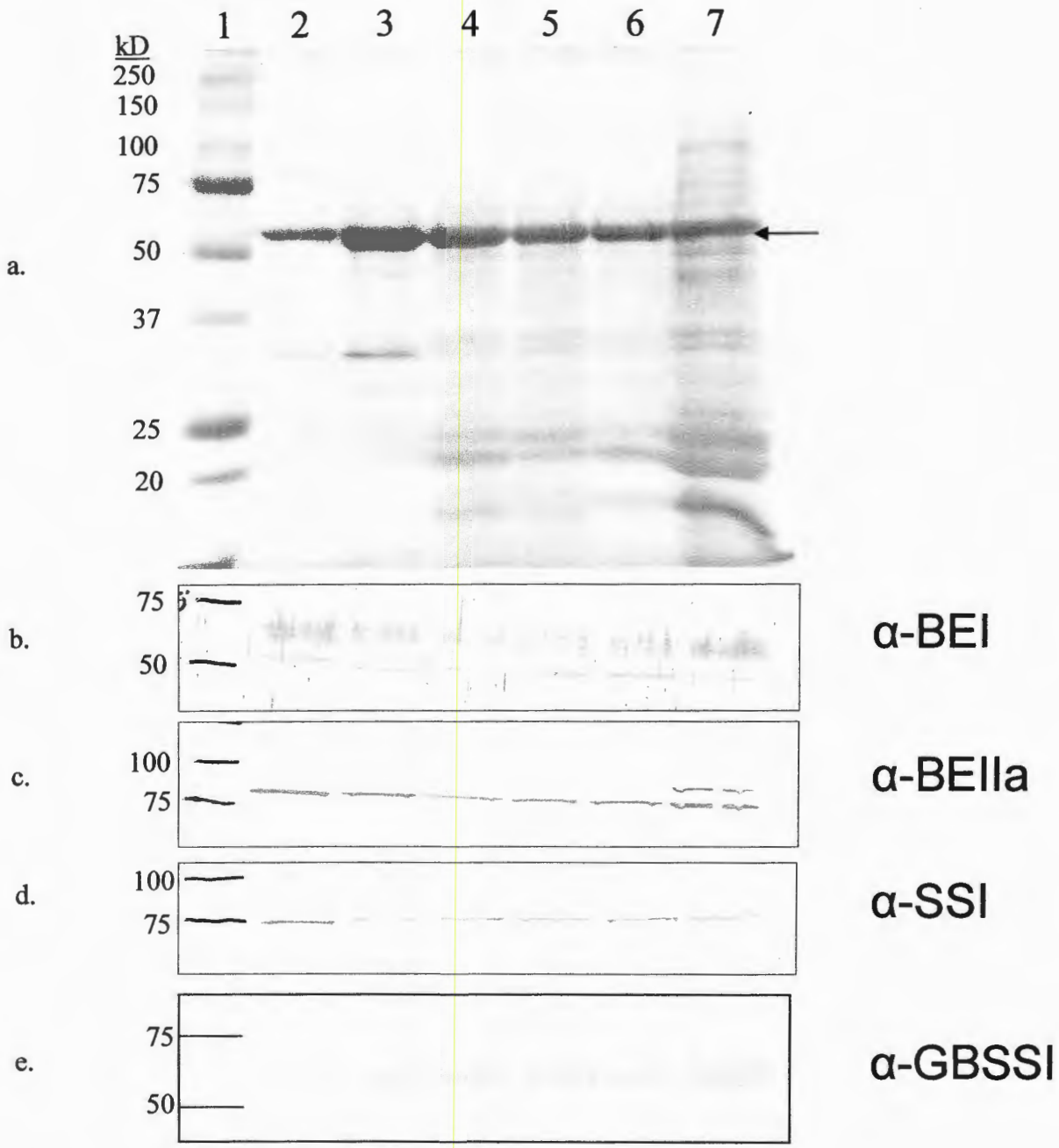


Figure 2: a.) Coomassie stain of thermolysin treated starch granules from wild type and *su2* mutants. b.) Immunodetection of thermolysin treated granules using α-BEI antibodies. c.) Immunodetection of thermolysin treated granules using α-BEIIa antibodies. d.) Immunodetection of thermolysin treated granules using α-SSI antibodies. e.) Immunodetection of thermolysin treated granules using α-GBSS. Lane assignments are as follows: 1-molecular weight markers, 2-wild type (W64A), 3-*su2*-2279, 4-*su2*-1982, 5-*su2*-19791, 6-*su2*-Ref, and 7-*su2*-5178. Only a slight lessening of band intensity is seen in SSI in *su2*-2279. The appearance of what looks like another protein around 60 kD in the mutants is indicated by the arrow.

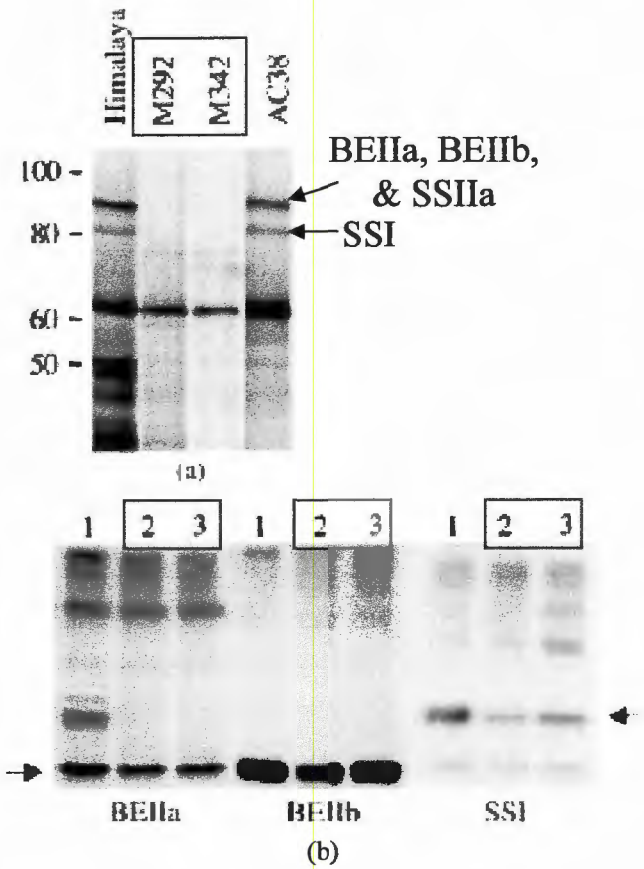


Figure 3: This figure is from Morell et al. [3] in order to compare the barley and maize experiments. Their description of the figure is as follows:
"Electrophoretic analysis of starch granule and soluble proteins from *sex6* mutants and wild type. All samples were from endosperm collected 15-18 days after anthesis. Molecular weight standards of defined mass (units are in kilodaltons) are indicated on the side of panel (a).
(a) 12.5% acrylamide (30: 0.135 Acryl:Bis), silver stained
(b) Non-denaturing PAGE analysis of soluble extracts of developing endosperm. Lanes marke 1 contain extracts of Himalaya, lanes marked 2 contain M292, and lanes marked 3 contain M342. Antibodies specific for BEIIa, BEIIb, and SSI were used as indicated beneath the panel. The arrow on the left side of this panel indicates the position at which both BEIIa and BEIIb migrate and the arrow on the right indicates SSI."

APPENDIX C: MAIZE DEHYDRINS

Introduction

Water deficit in plants can be induced by drought, high salinity, and low temperatures [1]. Plants will respond to this type of stress in many ways including alteration of gene expression, changes in metabolism, osmotic adjustment, induction of degradation and repair systems, and elevated expression of LEA (late embryogenesis abundant), chaperone, and mRNA-binding proteins [2]. However, plants with enhanced tolerance to these types of stresses have mostly been modified in a way that maintains the function and structure of cellular components [3]. Cereals are most sensitive to drought during two developmental stages: the onset of meiosis and in early grain initiation, also known as anthesis [4]. Dehydrins are LEA proteins that accumulate during dehydration and the focus of this research is to determine the level of dehydrin proteins in drought induced maize kernels at different developmental stages.

Dehydrins are proteins whose functions are hypothesized to help in maintaining the structure and function of the cell membrane and organelle membranes [5]. Two dehydrin proteins have been identified in maize embryo, DHN1 (18 kD) and DHN2 (40 kD), but other proteins are detected with dehydrin antibodies indicating that more are present but whose identities have yet to be confirmed. Dehydrins are in the family of LEA proteins in the D-11 group [6] which accumulate in plant tissues experiencing dehydration or by the application of abscisic acid (ABA). ABA is a phytohormone that plays a role in regulation of programmed dehydration in seeds and whose presence is sometimes induced by stress [7]. Dehydrins localize primarily in the nucleus, cytoplasm, and plasma membrane under stress conditions [8].

Dehydrins are characterized by a highly conserved region of 15 amino acids, EKKGIMDKIKEKLPG, which is known as the K segment. In dehydrins, this amino acid sequence can be repeated 1-11 times within a single protein [9]. The K segments form

amphipathic α-helices whose structures are similar to structures of other proteins including apolipoprotein's α-helices that occur in blood plasma lipoproteins which facilitate the transport of water-insoluble lipids in the plasma [10]. A suggested role of the K segment is that it may have hydrophobic interactions with partially denatured proteins or membranes [9].

The exact role that dehydrins play in water deficit stressed plants remains unknown. The hypothesis is that dehydrins are compatible solutes that act to stabilize macromolecules under stress which stabilizes the protoplasm [9]. The hydration of the hydrophobic regions, K segments, results "in the formation of an envelope of ordered water" to drive a partially unfolded protein back to a folded state [9]. This helps prevent injury to a cell, such as membrane destabilization, when it contracts due to water loss.

Drought tolerance is not completely understood and the research shows what effects drought has on maize seeds in different developmental stages following anthesis. Drought affects *Zea mays* in various ways, but the most obvious from our testing is smaller seed size. The hypothesis is that the seeds stop growing because of a transition to desiccation tolerance indicated by the accumulation of dehydrin protein. The results help explain why the seeds are smaller the earlier the plant is exposed to the stress based on the amount of dehydrin protein found in the seed.

Materials and Methods

Plants treatment and sampling

Plants were grown in a greenhouse under normal growing conditions in until anthesis occurred. There were four randomized treatments of drought stress: control (no water stress), blister (early stress, 15 days after anthesis), dough (mid stress, 25 days after anthesis), and first dent (late stress, 35 days after anthesis). Plant samples are approximately 5, 12, 20, 30, 38, 48, and 55 days after anthesis depending on stage of kernel development. Soil moisture was determined using a TDR probe of 5 inches in all pots every 2-3 days after 5 p.m. Kernel

water content and dry kernel weight was examined by the sampling of 10 kernels, acquisition of the fresh weight, and then drying at 80°C. Average weight of water loss and dry weight were determined from these 10 kernels. Approximately 40 kernels were froze in L N_2 and stored at -80°C which were used later for dehydrin level analysis.

Dehydrin analysis

Methods described for protein extraction were developed from Campbell *et al.* [11]. 100 g of each sample was homogenized in 600 µl of cold protein extraction buffer (Tris-MES or TES-HCl (pH = 7.5), 20 mM NaCl, and 1 mM PMSF) in an ice cold mortar and pestle. Homogenate was transferred to a 1.5 ml microfuge tube and stored on ice. Mortar and pestle was washed with an additional 600 µl of cold protein extraction buffer and was added to the microfuge tube. Then the samples were centrifuged at 10,000*g* for 5 min at 4°C. Supernatants were transferred to new tubes and were analyzed immediately or froze in L N_2 and stored at -80°C.

Proteins concentrations were determined using the Bradford assay [12] and loaded in equal amounts onto a 10% acrylamide Bio-Rad precast SDS-PAGE gel. Gels were run at 15 mA/gel through the stacker section and then at 20 mA/gel through the resolving section at room temperature until the dye front ran off. Proteins were detected via Coomassie stain or immunodetection using α-dehydrin antibodies received as a gift from Dr. Close, University of California, Riverside, Ca.

Results and Discussion

Four different treatments of stress were inflicted on the corn plants including: control (no water stress), blister (early stress, 15 days after anthesis), dough (mid stress, 25 days after anthesis), and first dent (late stress, 35 days after anthesis). To ensure proper development of the four treatments various physical conditions were analyzed. The first one evaluated the soil moisture content to show that less water is available in the stressed treatments compared

to the control (see Figure 1). Moisture content decreases rapidly once the water supply is cut off (varies with the treatment). The second evaluation of the treatments was to measure the water content in the kernels. The treatment also appears to effect the accumulation of water in the kernel (see Figure 2) indicating that drought stress has a direct effect on the kernel.

The earlier the onset of stress generally results in more severe damage to the plant. This is indicated by comparing the final dry weight of the kernels of all four treatments (see Figure 3). The accumulation of dry weight is highest in the control and first dent (late onset of stress), followed by dough, and then by blister. Remember that cereals are most sensitive to drought during two developmental stages: the onset of meiosis and anthesis [4]. The later the onset of drought stress from anthesis, the less damage is caused to the kernel which is evident in the dry weight accumulation data.

The next question would be why do the seeds stop growing? When water deficit occurs the plant will lose turbidity and the organellar membranes could become destabilized by coming in contact with one another. Dehydrins accumulate during the process of cellular dehydration to prevent this from happening. However, the entire cell is transitioned to desiccation tolerance because all of the macromolecules are stabilized by the dehydrin proteins. If the drought stress continues for a long period of time, the cell never reverts back to the normal cellular processes because of the dehydrin stabilization and further accumulation of dry weight is not possible. Figure 4 measures the amount of dehydrins accumulated in the four stress treatments. Dehydrin proteins accrue earlier in the blister and dough treatments compared to the control and 1st dent. This evidence supports the hypothesis is that the seeds stop growing because of a transition to desiccation tolerance indicated by the accumulation of dehydrin protein.

Further research is required to determine the exact function of the dehydrin protein in order to understand how it affects cellular processes. If the accumulation of dehydrin protein aids in drought tolerance perhaps this protein could be overexpressed in crops to prevent yield loss. Drought tolerance is not completely understood but this research shows that

drought has an effect on maize seeds in different developmental stages following anthesis based on decreased dry weight accumulation and earlier onset of accruing dehydrin proteins.

Acknowledgements

I would like to thank Octavio Caviglia for all of his work growing and testing the plants used in this experiment. I would also like to thank Rulian Zhou for her assistance in repeating a few of the tests to ensure we had good quality results.

References

1. Buchanan, B.B., W. Gruissem, and R.L. Jones, *Biochemistry & Molecular Biology of Plants*. 2000, Rockville, Maryland: American Society of Plant Physiologists. 1367.
2. Koag, M.C., et al., *The binding of maize DHN1 to lipid vesicles. Gain of structure and lipid specificity.* Plant Physiol, 2003. **131**(1): p. 309-16.
3. Wang, W., B. Vinocur, and A. Altman, *Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance.* Planta, 2003. **218**(1): p. 1-14.
4. Saini, H.S. and M.E. Westgate, *Reproductive development in grain crops during drought.* Advances in Agronomy, 2000. **68**: p. 59-96.
5. Wise, M.J. and A. Tunnacliffe, *POPP the question: what do LEA proteins do?* Trends Plant Sci, 2004. **9**(1): p. 13-7.
6. Dure, L., 3rd, *A repeating 11-mer amino acid motif and plant desiccation.* Plant J, 1993. **3**(3): p. 363-9.
7. Allagulova Ch, R., et al., *The plant dehydrins: structure and putative functions.* Biochemistry (Mosc), 2003. **68**(9): p. 945-51.
8. Borovskii, G.B., et al., *Accumulation of dehydrin-like proteins in the mitochondria of cereals in response to cold, freezing, drought and ABA treatment.* BMC Plant Biol, 2002. **2**(1): p. 5.

9. Close, T.J., *Dehydrins: emergence of a biochemical role of a family of plant dehydration proteins.* Physiologia Plantarium, 1996. **97**: p. 795-803.

10. Campbell, S.A. and T.J. Close, *Dehydrins: genes, proteins, and associations with phenotypic traits.* New Phytol, 1997. **137**: p. 61-74.

11. Campbell, S.A., et al., *A ca. 40 kDa maize (Zea mays L.) embryo dehydrin is encoded by the dhn2 locus on chromosome 9.* Plant Mol Biol, 1998. **38**(3): p. 417-23.

12. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.

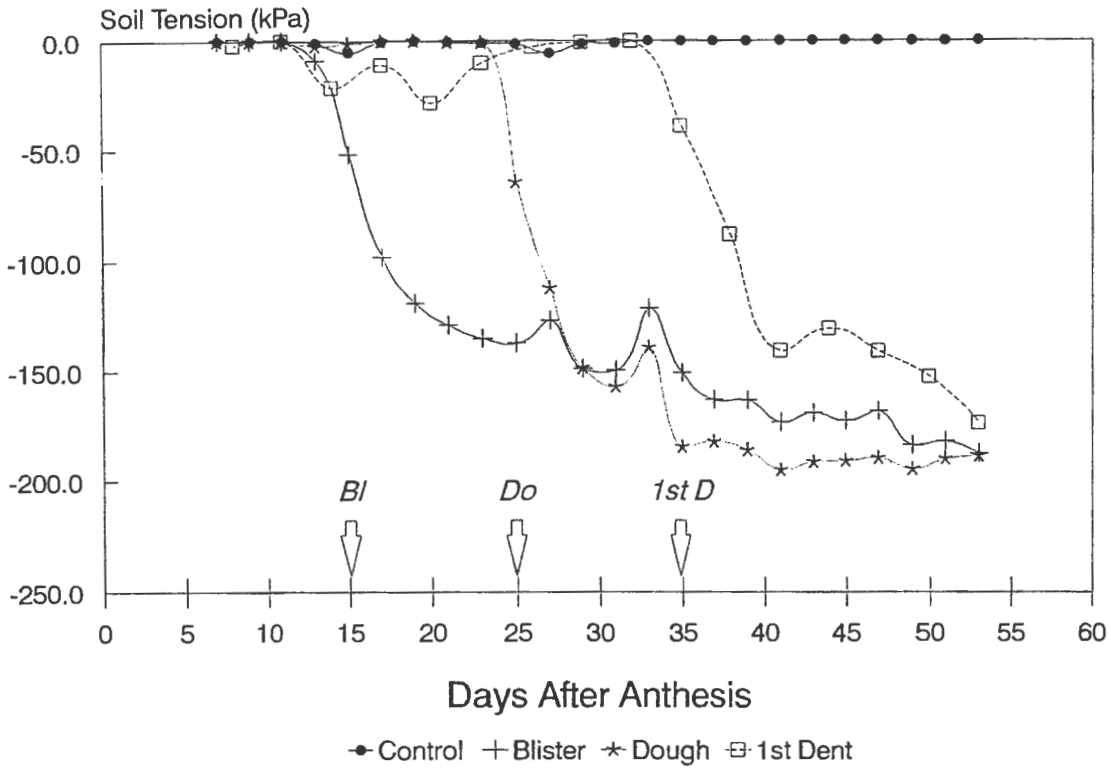


Figure 1: Soil moisture tension of early cut-off treatments

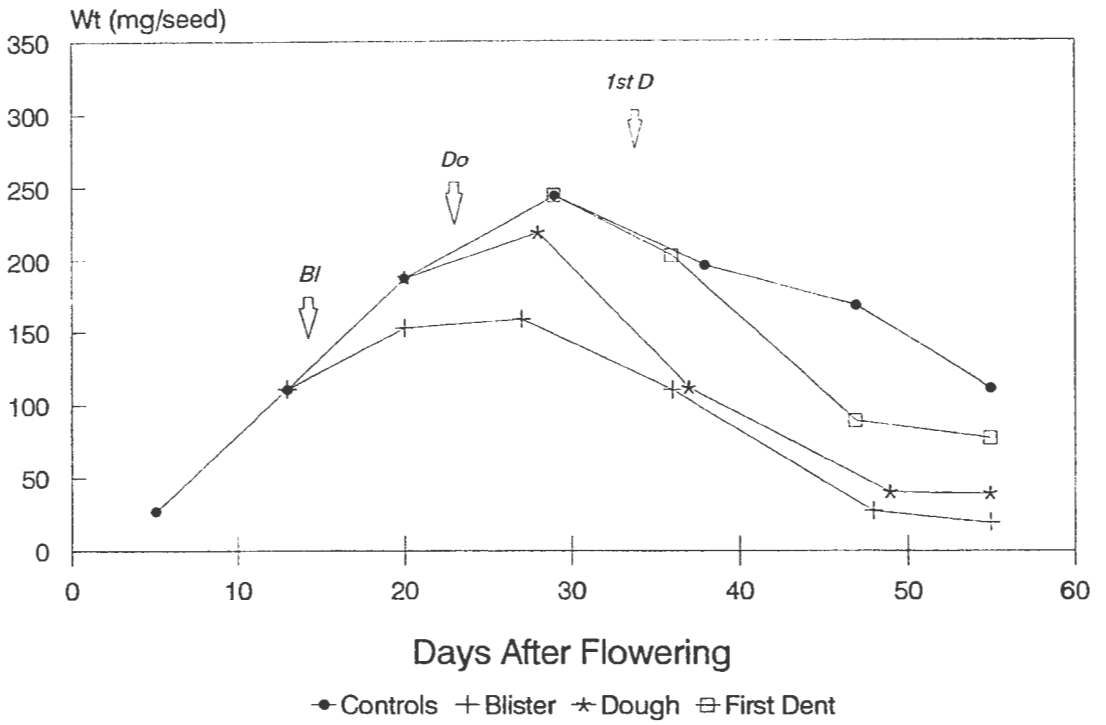


Figure 2: Kernel water content of early cut-off treatments

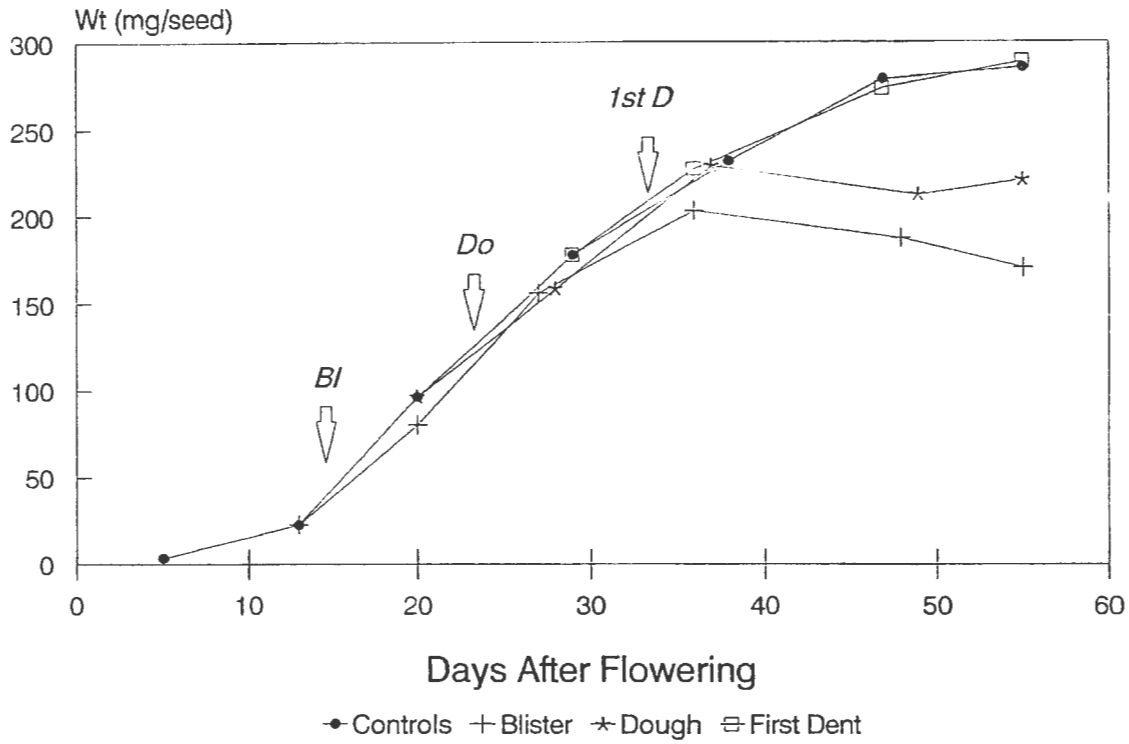


Figure 3: Kernel dry weight accumulation of early cut-off treatments

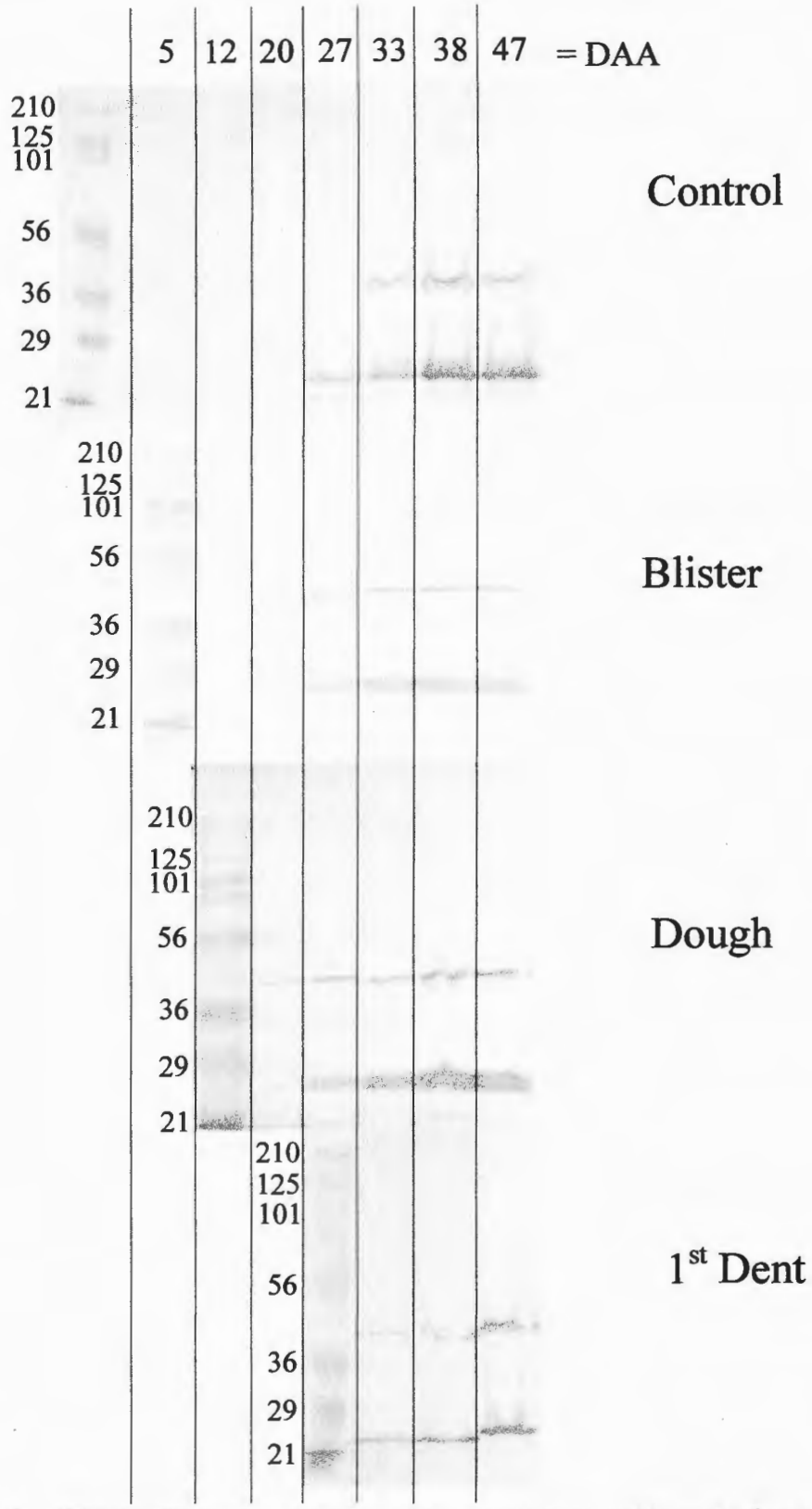


Figure 4: SDS-PAGE immunodetection with α-dehydrin (1:1000) on the four stress treatments.